

**INDIAN NATIONAL  
SCIENCE ACADEMY  
AWARD LECTURES  
(1984-1993)**



**A DIAMOND JUBILEE PUBLICATION  
(1994)**



**INDIAN NATIONAL SCIENCE ACADEMY**  
Bahadur Shah Zafar Marg, New Delhi-110002

# ***Indian National Science Academy Award Lectures***

**A Compilation of Academy  
Award Lectures published in  
the Journals of the Academy,  
representing the significant  
achievements in Science and  
Technology during the last  
decade.**



**INDIAN NATIONAL  
SCIENCE ACADEMY  
AWARD LECTURES  
(1984-1993)**

**(Volume IV)**

**DIAMOND JUBILEE PUBLICATION**



**INDIAN NATIONAL SCIENCE ACADEMY  
Bahadur Shah Zafar Marg, New Delhi-110002**

© Indian National Science Academy

Published in January 1994

Price Rs. 300 US \$ 100

Published by S K Sahu, Executive Secretary, Indian National Science Academy, Bahadur Shah Zafar Marg, New Delhi-110002 and Printed at Aakriti Graphics, B-62/10, Naraina Industrial Area, Phase-II, New Delhi-110028 ©: 5700089

## F O R E W O R D

The present publication of the Academy Award Lectures is a useful effort of the Academy to make available to a wide audience the best achievements of its Fellows. An Award Lecture would reveal the best concerted effort of a scientist and it is good that INSA thought of bringing together most of these Academy Award Lectures together in this form.

One of the activities of the Academy is to honour and recognize excellence in science and technology. Academy has, thus, instituted 50 medals/lectures embracing almost all aspects of science & technology. Some of the medals have been instituted by the Academy out of its own funds and others by the endowments received from donors. The awards come under categories international, general and subject-wise awards.

Academy awards are conferred upon the recipients annually after he/she delivers the award lecture either at the General Meeting, Anniversary General Meeting or under the aegis of a Local Chapter of the Academy in accordance with the regulations governing the awards.

During the Golden Jubilee of the Academy, it was decided to compile all the award lectures which have been delivered by the recipients ever since their inception till 1983. These award lectures were brought out in two volumes and this endeavour was highly appreciated by the scientific community. A large number of these lectures have laid the foundation for the future of science & technology in this country and also reflected the progress and development of Indian Science during this period. Keeping in mind the overwhelming response to the previous two volumes, I am happy to place before you volumes III & IV of this series which include the award lectures delivered from the year 1983 to 1993. Out of a total of 142 lectures which have been delivered during this period, volume III includes 43 lectures in physical sciences and volume IV 54 lectures in basic and biological sciences as well as those relating to general medals and/or lectures of the academy. A few award lectures are unfortunately not available with the Academy and as such, despite our best efforts, could not be included in these volumes.

I am confident that volumes III & IV of INSA Award Lectures will be a useful document for the scientific community.

I wish to thank Professor J.C. Ahluwalia and staff of the Academy for their dedicated efforts to bring out these volumes.

January 7, 1994  
New Delhi

**S.K. Joshi**  
President, INSA



## P R E F A C E

In the present world of complex communications and high technical knowledge, the “artist” has become the scientist and the “scientist” has become the aesthetic artist. The novelty of science and technology has become continuous. The artist is able to “paint” through pixels colourfully an endless expression of his truths. Such is the merger of scientist into an artist, which if unrecognized by the society as a matter of human catharsis, may destroy the most beautiful creative genius in a scientist. Thus, there is all the need for “Awards” which remain the most effective tools of preservation and promotion of natural knowledge.

As a part of its “objectives”, the Academy has instituted fifty awards under categories- (a) international, (b) general medals and lectures (c) subject-wise medals and (d) young scientist awards. A brief general introduction to these coveted Academy Awards is presented here.

## AWARDS AT A GLANCE

### INTERNATIONAL AWARDS

#### (i) *INSA-Vainu Bappu Memorial Award*

Established in 1985 from an endowment of Rs. 3 lakhs by Mrs Sunanna Bappu, mother of the late Dr Manali Kallat Vainu Bappu, an eminent Astronomer and a Fellow of the Academy, the award (carrying a financial value of Rs. 25,000/-) is given annually to an Astronomer/Astrophysicist of international recognition. The first award was announced in 1985. The recipient is also given a bronze medal.

The presentation of the Award is usually made to the recipient at the Anniversary Meeting or at a subsequent General Meeting at which he/she delivers his/her lecture.

#### (ii) *Jawaharlal Nehru Birth Centenary Medal*

Established in 1989, the award is made once in three years for international cooperation in science & technology. The first award was made in 1990. The award carries a bronze medal and a citation.

The presentation of the award is made at the time when the recipient delivers the award lecture.

## **GENERAL MEDALS AND LECTURE AWARDS**

- **The Chandrasekhara Venkata Raman Medal**
- **The Shanti Swarup Bhatnagar Medal**
- **The Meghnad Saha Medal**
- **The Aryabhata Medal**

The above Medals shall be awarded every two years for outstanding contributions in any branch of science coming within the purview of the Academy.

Two out of the four General Medals will be awarded each year, one in biological sciences and the other in physical sciences. Thus, in the year 1991, *C V Raman Medal* and *S S Bhatnagar Medal* were awarded and in 1992 *Meghnad Saha Medal* and *Arybhata Medal* were awarded.

The presentation of the medal is made to the recipient at the Annual General Meeting of the year of the award or at the General Meeting at which he delivers his lecture.

The award consists of a Copper medal (gold plated).

## **Indira Gandhi Prize for Popularization of Science**

The award was instituted by INSA in 1986 to encourage and recognize popularization of science.

The prize is awarded once in two years for outstanding work done by an individual for the popularisation of science in any Indian language, including English. The awardee is a distinguished writer, editor, journalist, lecturer, radio or television programme director, science photographer or an illustrator, which has enabled him/her to interpret science (including medicine), research and technology to the public. The awardee is also expected to have a knowledge of the role of science, technology and research in the enrichment of cultural heritage and in solution of problems of humanity. The first award was made in 1986.

The prize is open to any Indian national residing in the country and carries a cash award of Rs. 10,000/- and a bronze medal.

The awardee is judged from his writings in newspapers, magazines, popular books and scripts prepared for radio and television programmes and films. Initiative, originality, scientific accuracy, clarity of interpretation and impact

of creating excitement, interest and understanding of science are the important criteria for judging the entries.

## **LECTURES**

### **i) The Sisir Kumar Mitra Memorial Lecture**

The lecture was established in the year 1963 by the Academy from the General Funds of the Academy in the memory of the late Professor Sisir Kumar Mitra, a Foundation Fellow of the Academy. The first award was made in the year 1966.

### **ii) The Kariamanikkam Srinivasa Krishnan Memorial Lecture**

The Lecture was established in the year 1965 by the Academy from the General Funds of the Academy in the memory of the late Professor Kariamanikkam Srinivasa Krishnan, a Foundation Fellow of the Academy. The first award was made in the year 1969.

The presentation of the award is usually made to the recipient at the Annual General Meeting or at a subsequent General Meeting at which he/she delivers his/her lecture.

### **iii) Jawaharlal Nehru Birth Centenary Lecture**

Established in 1989 from the general funds of the Academy, two lectureships are awarded every year to Indian scientists. The Lecture carries a cash award of Rs. 5,000/-. The first award was made in the year 1990.

The presentation of the awards is made at the time when the recipients deliver the award lecture.

## **SUBJECTWISE MEDALS/LECTURES/AWARDS**

### **a) Medals Instituted by the Academy**

#### **(i) *The Srinivasa Ramanujan Medal***

— for Mathematics or a related subject

#### **(ii) *The Satyendranath Bose Medal***

— for Theoretical Physics

- (iii) ***The Homi Jehangir Bhabha Medal***  
— for Experimental Physics
- (iv) ***The Jagadish Chandra Bose Medal***  
— Biochemistry, Biophysics, Molecular Biology & related areas
- (v) ***The Sunder Lal Hora Medal***  
— for Plant and Animal Sciences
- (vi) ***The Darashaw Noshervanji Wadia Medal***  
— for Earth Sciences, Geology, Geophysics & Geography
- (vii) ***The Prasantha Chandra Mahalanobis Medal***  
— for Engineering & Technology
- (viii) ***The Syed Husain Zaheer Medal***  
— for Engineering & Technology
- (ix) ***The Silver Jubilee Commemoration Medal***  
— for Agriculture and Applied Sciences
- (x) ***The Golden Jubilee Commemoration Medal***  
— for Chemical Sciences
- (xi) ***The Golden Jubilee Commemoration Medal***  
— for Biological Sciences
- (xii) ***The Kalpathi Ramakrishna Ramanathan Medal***  
— for Atmospheric Sciences and Meteorology

The medals are awarded every three years for outstanding contributions to the discipline of science concerned. Eminence is judged in accordance with the criterion that the scientific work of the candidate is such that its impact has been felt for a considerable length of time.

The person to whom the award is made would deliver a lecture on the subject of his work at the Anniversary General Meeting in January of the year to which the award relates or at one of the General Meetings of the Academy.

## **b) Endowed Medals**

### **1. Vishwakarma Medal**

Established in 1976 from an endowment of Rs. 30,000/- by Dr Pulin Bihari Sarkar, FNA, the Medal is awarded to eminent scientists, technologists or



any one whose discovery or invention has led to the start of a new industry in India or to a significant improvement of an existing process resulting in cheaper or better product. The medal carrying a bronze medal and cash award of Rs. 7,500/- is to be given every three years. The first medal was awarded in 1979.

**2. *Professor Brahm Prakash Memorial Award***

Established in 1987 from an endowment from Mrs R Brahm Prakash and subsequently donated by friends of late Dr Brahm Prakash, an eminent Fellow of the Academy, the award (carrying a cash of Rs. 10,000/- and a bronze medal) is given every third year to a scientist/engineer who has made outstanding contributions in any area of Engineering and Technology.

**3. *Professor Shyam Bahadur Saksena Memorial Award***

Established in 1989 from an endowment of Rs. 50,000/- by Mrs Sarla Saksena to commemorate the memory of late Professor Shyam Bahadur Saksena, a distinguished botanist and Fellow of the Academy, the award (carrying an amount of Rs. 10,000/- and a bronze medal) is given every three years to an eminent scientist who has made outstanding contributions in any branch of Botany. The first award was made in 1990.

**4. *Professor G N Ramachandran 60th Birthday Commemoration Medal***

Established in 1989 out of an endowment of Rs. 1 lakh by the Organizing Committee to commemorate the 60th Birthday of Professor G N Ramachandran, a distinguished Fellow of the Academy the Award (carrying a cash prize of Rs. 20,000 and a bronze medal) is given once in three years to an eminent scientist for outstanding contributions in the field of Molecular Biology, Biophysics and Crystallography. The recipient delivers a lecture on the subject of his work at the Anniversary General Meeting in January of the year to which the award relates or at any of the General Meetings of the Academy.

**5. *Professor Tiruvenkata Rajendra Seshadri Seventieth Birthday Commemoration Medal***

Established in 1971 from an endowment of Rs. 10,000/- by students of Professor Seshadri, an eminent organic chemist and a Fellow of the Academy, the Medal (carrying an amount of Rs. 5,000/- and a bronze medal) is awarded every three years to an eminent chemist of Indian nationality for outstanding work in any branch of chemistry or chemical technology. The first medal was awarded in 1973. The recipient delivers a lecture on a subject of his choice under the auspices of a Local Chapter.

## 6. *Professor K P Bhargava Memorial Medal*

The lecture was recently established (in 1992) out of an endowment of Rs. 50,000/- made by Mrs Savitri Bhargava, to commemorate the memory of late Professor K P Bhargava, an eminent pharmacologist and a distinguished Fellow of the Academy.

The selection of the scientist is to be made from amongst the persons, below the age of 50 years who have made outstanding contributions in the area of Basic Medical Sciences. The lecture award is to be given once in three years. The first award will be made in 1996. The award carries an honorarium of Rs. 10,000/- plus TD/DA for the journey performed to deliver the award lecture. The recipient shall deliver the lecture of his/her choice under the aegis of a Local Chapter.

7. *Chandrakala Hora Memorial Medal* (established in 1945) from an endowment of Rs. 3,000/- by Dr Sunder Lal Hora and Mrs Hora in memory of their daughter, is awarded (every fifth year) to an eminent scientist for conspicuously important contributions to the development of fisheries, aquatic biology and related areas in India during the five years preceding the year of award. The medal is to be given every fifth year. The first medal was awarded in 1950.
8. *Shri Dhanwantari Prize* (established in 1969) from an endowment of Rs. 16,000/- + Rs. 2,500/- by Shri Anant Krishna Asundi in the memory of his youngest daughter Shrimati Akkadevi, is awarded (every fifth year) to eminent scientists for outstanding work in India in any branch of medical sciences in its widest sense, including research in drugs and methodology of Ayurveda. This includes research in medical as well as chemical, physical and biological sciences aimed at the amelioration of human suffering. In fact, its scope shall include any outstanding discovery in drug or mode of treatment or inventions established as a landmark in medical science in its widest sense. The first prize was given in 1971.

## c) **Endowment Lectures**

1. *Professor Bal Dattatraya Tilak Lecture Award* established (in 1982) out of an endowment of Rs. 1,00,000 by Professor B D Tilak Scientific Research and Education Trust, C/O National Chemical Laboratory, Pune, to commemorate Professor B D Tilak, a Fellow of the Academy distinguished for his researches in the field of dyestuffs chemistry and organic chemical technology, is made (every year) to a person for outstanding contributions to rural economy and life through innovative and effective application of science and technology. The first award was made in 1983.

2. ***Dr T S Tirumurti Memorial Lecture*** established (in 1985) out of an endowment of Rs. 25,000/- by Mrs Janaki Ramachandran, daughter of late Dr Tirunelveli Subbaiyer Tirumurti, a Foundation Fellow of the Academy, who had made notable contributions to pathology and medicine, is made (every two years) to a person for outstanding contributions in the field of Medical Sciences. The first award was made in 1985.
3. ***Dr Nitya Anand Endowment Lecture*** established (in 1986) out of an endowment of Rs. 1,30,000/- by the organizing committee to celebrate the 60th Birthday of Dr Nitya Anand, an eminent chemist and a Fellow of the Academy, is delivered every alternate year by a scientist below 50 years of age who has done outstanding work during the previous decade in the area of biomedical research including new drug development.
4. ***INSA Prize for Materials Science*** established (in 1986) out of the Endowment of Rs. 50,000/- by the Organizing Committee of the International Conference on the 'Application of Mössbauer Effect', held in 1981, is awarded every alternate year for outstanding contributions in Materials Science. The first award was made in 1987.
5. ***Professor S Swaminathan 60th Birthday Commemoration Lecture Award*** established (in 1990) out of an endowment of Rs. 75,000 by Professor S Swaminathan 60th Birthday Commemoration Committee to commemorate Professor S Swaminathan, a Fellow of the Academy, distinguished for his research in the field of Organic Chemistry (Synthetic and Applied).
6. ***Professor M R N Prasad Memorial Lecture Award*** established (in 1989) out of an endowment of Rs. 50,000/- by the colleagues and friends of late Professor M R N Prasad, FNA, to commemorate his memory, is made (every three years) to a person for outstanding contributions in the field of Animal Physiology in its widest sense.
7. ***The Bires Chandra Guha Memorial Lecture*** is made (every three years) to a person for distinct contributions in the field of Biochemistry, Nutrition, Food and allied problems in the broadest sense. The first award was made in 1969.
8. ***The Bashambar Nath Chopra Lecture*** established (in 1968) out of an endowment of Rs. 10,000/- in memory of Dr Bashambar Nath Chopra, FNA by his family, is made (every third year) to a scientist from amongst the persons who have made distinct contributions to any branch of biological sciences coming within the purview of the Academy. The first award was made in 1971.

9. ***Professor Shambu Nath De Memorial Lecture Award*** was established in 1992 out of an endowment of Rs. 50,000/- by the Organizing Committee of the VII International Specialised Symposium on Yeasts (IUMS) on behalf of the Association of Microbiologists of India (AMI) to commemorate Professor S N De, distinguished scientist for his research in Microbiology.

The selection of the scientist is to be made from amongst the persons who have made outstanding contributions in the area of microbiology in the widest sense. The lecture award is to be given once in three years. The first award has been announced in 1993.

10. ***Professor Vishnu Vasudeva Narlikar Memorial Lecture Award*** was established in 1992 out of an endowment of Rs. 1 lakh by the family members and a close friend of late Professor V V Narlikar, to commemorate Professor V V Narlikar, an eminent Mathematician and a Fellow of the Academy.

The selection of the scientist is to be made from amongst the persons who have made outstanding contributions in the field of Applied Mathematics including Gravitational Theory. The lecture award is to be given once in three years. The first award will be made in 1994.

11. ***Dr Jagdish Shankar Memorial Lecture*** was established in 1992 to commemorate the memory of Dr Jagdish Shankar, a Fellow of the Academy and a distinguished scientist in the field of Nuclear and Radiation Chemistry out of an endowment of Rs. 58,534/- by the Organising Committee of the Emerging Frontiers in Chemistry Symposium, students, admirers and family members of Dr Jagdish Shankar.

The selection of the scientist is to be made from amongst the persons, below the age of 50 years, who have made outstanding contributions in the area of Chemical Sciences in its widest sense. The lecture award is to be given once in three years. The first award will be made in 1994.

12. ***Professor K Rangadhama Rao Memorial Lecture*** established (in 1979) out of an endowment of Rs. 10,000/- by old students of Professor K Rangadhama Rao, an eminent physicist and a distinguished Fellow of the Academy, is made to a person once in four years known for his outstanding contributions in the field of Spectroscopy. The lecture is to be given once in four years alternately with Professor R K Asundi Memorial Lecture. The first award was made in 1979.

13. ***Professor Rango Krishna Asundi Memorial Lecture*** established (in 1983) out of an endowment of Rs. 21,000/- by the Asundi Endowment Fund to commemorate Professor Rango Krishna Asundi, a Fellow of the Academy distinguished for his researches in spectroscopy, is made to person

(once in four years) for outstanding contributions in the field of Spectroscopy. The lecture is given once in four years alternately with Professor K Rangadhama Rao Memorial Lecture. The first award was made in 1984.

14. ***Professor Toppur Seethapathi Sadasivan Lecture Award*** established (in 1982) out of an endowment of Rs. 25,000/- by the Professor Toppur Seethapathi Sadasivan Endowment Committee to commemorate Professor Sadasivan, a Fellow of the Academy distinguished for his researches in physiology and plant pathology, is made to a person for outstanding contributions in any field of Botany. The lecture is given once in four years alternately with Professor P Maheshwari Memorial Lecture. The first award was made in 1984.
15. ***Professor Panchanan Maheshwari Memorial Lecture*** established (in 1984) out of an endowment of Rs. 26,200/- by the colleagues and friends of late Professor Panchanan Maheshwari, a distinguished botanist and a Fellow of the Academy, is made to a person who had made outstanding contributions in any area of Plant Sciences
16. ***Dr. Guru Prasad Chatterjee Memorial Lecture*** established (in 1979) out of an endowment of Rs. 15,000/- by Dr G P Chatterjee and Mrs Suniti Chatterjee to commemorate Dr Guru Prasad Chatterjee, an eminent metallurgist and Fellow of the Academy, is made (once in five years) to a scientist known for his outstanding contributions in any field of science as recognized by INSA. The first award was made in 1981.
17. ***Dr Har Swarup Memorial Lecture*** established (in 1981) out of an endowment of Rs. 10,000/- by Dr (Mrs) Savitri Swarup to commemorate Dr Har Swarup, a Fellow of the Academy distinguished for his researches on endocrinology, physiology and developmental biology, is made (once in five years) to persons who have made outstanding contributions in the field of zoology. The first award was made in 1984.

Further there are three young scientist awards — 1) INSA Medal for Young Scientists, 2) Professor LSS Kumar and Anil Kumar Bose Memorial awards to encourage potential of scientific research in younger generation. On an average 15 awards are made every year.

Certain essential features contribute to procedure for selection and other conditions relating to the Academy awards like e.g., 1. Notice of the Award 2. Scrutiny of Nominations 3. Appointment of the Advisory Board 4. Circulation of the List to Members of the Advisory Board 5. Selection by the Advisory Board 6. Selected Name to be Communicated to the Council 7. Announcement 8. Presentation of award and delivery of lecture 9. Specification of the Medal etc.

I am sure that these volumes will be liked by everyone among the scientific community as well as other readers. I appreciate the help and dedicated efforts of the staff of the Academy in bringing out these two volumes.

January, 1994  
New Delhi

***J.C. Ahluwalia***  
Secretary, INSA

# C O N T E N T S

|   | Page No   |
|---|-----------|
| Foreword  | (iii)     |
| Preface   | (v)       |
| <b>C V RAMAN MEDAL</b>  |           |
| 1. C Gopalan  | 999-1022  |
| 2. A P Mitra  | 1023-1065 |
| <b>S S BHATNAGAR MEDAL</b>                                      |           |
| 3. Dinesh Mohan   | 1066-1084 |
| 4. B K Bachhawat  | 1085-1099 |
| <b>MEGH NAD SAHA MEDAL</b>                                      |           |
| 5 R Ramanna   | 1100-1139 |
| 6 Sukh Dev  | 1140-1177 |
| 7 C N R Rao   | 1178-1195 |
| 8 S Chandrasekhar   | 1196-1214 |
| <b>ARYABHATA MEDAL</b>  |           |
| 9. P N Srivastava   | 1215-1227 |
| 10. O Siddiqi   | 1228-1236 |
| <b>JAWAHARLAL NEHRU BIRTH CENTENARY LECTURE</b>                 |           |
| 11. S S Jha   | 1237-1254 |
| 12. M K Chandrasekharan   | 1255-1268 |
| 13. N Mukunda   | 1269-1285 |
| 14. Bimla Buti  | 1286-1301 |
| 15. K K Kannan  | 1302-1320 |
| <b>JAGADISH CHANDRA BOSE MEDAL</b>                              |           |
| 16 V Sasisekharan   | 1321-1336 |
| 17 N Appaji Rao   | 1337-1352 |
| 18 P K Maitra   | 1353-1361 |
| <b>SUNDER LAL HORA MEDAL</b>                                    |           |
| 19 P N Mehra  | 1362-1401 |
| 20. C M S Dass  | 1402-1409 |
| <b>SILVER JUBILEE COMMEMORATION MEDAL</b>                       |           |
| 21. S K Mukherjee   | 1401-1422 |
| <b>GOLDEN JUBILEE COMMEMORATION MEDAL (Biological Sciences)</b> |           |
| 22 M S Kanungo  | 1423-1438 |
| <b>SHYAM BAHADUR SAXENA MEDAL AWARD</b>                         |           |
| 23. H Y Mohan Ram   | 1439-1464 |
| <b>CHANDRAKALA HORA MEMORIAL AWARD</b>                          |           |
| 24. N B K Nair  | 1465-1498 |
| <b>SHREE DHANWANTRI PRIZE</b>                                   |           |
| 25. P N Tandon  | 1499-1518 |
| 26. M V S Valiathan   | 1519-1526 |

## B D TILAK LECTURE

|     |              |           |
|-----|--------------|-----------|
| 27  | A K N Reddy  | 1527-1544 |
| 28  | C V Seshadri | 1545-1559 |
| 29  | S S Kalbag   | 1560-1566 |
| 30. | P K Sethi    | 1567-1583 |
| 31  | V Kurien     | 1584-1595 |
| 32  | N H Antia    | 1596-1613 |

## T S TIRUMURTI MEMORIAL LECTURE

|    |             |           |
|----|-------------|-----------|
| 33 | B Ramamurti | 1614-1527 |
| 34 | S C Seal    | 1628-1671 |

## NITYA ANAND ENDOWMENT LECTURE

|    |                |           |
|----|----------------|-----------|
| 35 | Indira Nath    | 1672-1683 |
| 36 | C M Gupta      | 1684-1897 |
| 37 | B S Srivastava | 1698-1706 |

## BIRES CHANDRA GUHA MEMORIAL LECTURE

|    |                |           |
|----|----------------|-----------|
| 38 | B K Bachhawat  | 1707-1719 |
| 39 | Mahtab S Bamji | 1720-1731 |
| 40 | N K Notani     | 1732-1740 |

## BASHAMBAR NATH CHOPRA MEMORIAL LECTURE

|    |              |           |
|----|--------------|-----------|
| 41 | P M Bhargava | 1741-1746 |
|----|--------------|-----------|

## PANCHANAN MAHESHWARI MEMORIAL LECTURE

|    |                 |           |
|----|-----------------|-----------|
| 42 | T V Desikachari | 1747-1760 |
| 43 | V Puri          | 1761-1789 |
| 44 | D D Awasthi     | 1790-1803 |

## T S SADASIVAN LECTURE

|    |                 |           |
|----|-----------------|-----------|
| 44 | R Mishra        | 1804-1810 |
| 45 | B M Jauhari     | 1811-1827 |
| 46 | C V Subramaniam | 1828-1845 |
| 47 | D D Pant        | 1846-1873 |

## G P CHATTERJEE MEMORIAL LECTURE

|     |             |           |
|-----|-------------|-----------|
| 48  | B N Singh   | 1874-1894 |
| 49. | T N Khoshoo | 1895-1907 |
| 50. | Prem Narain | 1908-1931 |
| 51  | R Rajaraman | 1932-1942 |

## HAR SWARUP MEMORIAL LECTURE

|     |                |           |
|-----|----------------|-----------|
| 52. | M L Roonwal    | 1943-1983 |
| 53. | Ishwar Prakash | 1984-1191 |

## AUTHOR INDEX

1992





**Coluthur Gopalan** (b. 29 November 1918) did M.D (1945) from University of Madras, Ph D (1949) and D.Sc. from University of London, UK, FRCP, University of Edinburgh, University of London, UK, D.Sc. (h.c.), Banaras Hindu University, Varanasi. He is President, Nutrition Foundation of India and was formerly Director-General, Indian Council of Medical Research, New Delhi (1974-79), Director, National Institute of Nutrition, Hyderabad (1960-74)

Gopalan's researches have brought to fore various aspects of nutrition including public health, agriculture and socio-economic aspects, especially with regard to underprivileged communities. Time and again he has sought to underscore the importance of nutrition as a major factor in national and human development.

Gopalan is Fellow of Royal Society of London, Indian Academy of Sciences, and National Academy of Medical Sciences (India), Honorary Member, American Institute of Nutrition, President, Nutrition Society of India, President, Nutrition Foundation of India. He was Member, INSA Council (1969-71). He is the recipient of Basanti Devi Amir Chand Prize (1954) and Basanti Devi Amir Chand Prize (Senior) (ICMR) (1960), Amrut Mody Research Award (1972); Dr B C Roy National Award (1974), Ambhuj Nath Bose Prize (Royal College of Physicians, London) (1975), Ademola Prize (London School of Hygiene and Tropical Medicine) (1976), Dhanwantari Award (1978); FICCI Award (1978), WHO 'Health for All' Medal (1988), International Union of Nutrition Sciences (IUNS) Award (1989), C V Raman Medal (INSA) (1988), R D. Birla Award (1990)

*Coluthur Gopalan was elected to the fellowship of the Academy in 1966*

**Address:** President, Nutrition Foundation of India, B-37, Gulmohar Park, New Delhi 110 049, Tel (011) 4619431/319, Res Tel: (011) 669254.

## **COMBATING MALNUTRITION—PRACTICAL LEADS FROM SCIENTIFIC RESEARCH IN INDIA**

C GOPALAN FNA

The term 'malnutrition', in this presentation, is being used in a restricted sense to refer to major health disorders in the country associated with undernutrition and poverty. We are not discussing here the disorders of the affluent generally associated with 'malnutrition'—of a different kind namely 'overnutrition', though these are not unimportant. Undernutrition, however, is today the single most important factor undermining health, productivity and quality of our human resources and must, therefore, command immediate major concern and highest priority. Undernutrition and the state of ill health associated with it arise not just from dietary deficiency alone but from the synergistic interaction of poor dietaries and poor environmental hygiene—the two major attributes of the pervasive "poverty syndrome".

It is, no doubt, true that, in the ultimate analysis, the eradication of undernutrition can only be achieved through socioeconomic development and elimination of poverty. But this consideration should in no way obscure the overriding need for scientific research in Human Nutrition and the important practical contributions that such research can offer. In India, as in all poor developing countries, there is a whole spectrum of specific nutritional deficiency diseases, the precise pathogenesis and approach to prevention of each of which needs careful elucidation. In situations, like those obtaining in most developing countries, where undernutrition is widespread and resources to combat it are scarce, scientific research in nutrition acquires special relevance and importance and can help to identify feasible and cost-effective strategies for prevention and control. Health disorders associated with affluence and overnutrition, like obesity and atherosclerosis, for example, mostly afflict populations of prosperous developed countries and, therefore, naturally receive major attention and research support in those countries. On the other hand, scientific research on diseases of undernutrition must necessarily have to find top priority in the scientific agenda of poor developing countries. Undernutrition is *our* problem, not *theirs*. However, this is not to deny that outstanding basic contributions to nutrition science, which have benefited mankind in

general (including developing and developed countries) have emerged from research in developed countries.

The major point that will be made in this presentation is that scientific research in India on problems of undernutrition during the last three decades has in fact helped to generate several practical leads for action. It may be legitimately claimed that some of these contributions have been pioneering and have helped to enrich Nutrition Science in general. The sad fact that the several practical leads thus generated through painstaking research, have not always been put to optimal use because of inadequacies in our Health and Rural Development systems cannot detract from their inherent scientific merit and practical value. Some *examples* of practical leads that had emerged from scientific research in Human Nutrition in the country in recent years are briefly discussed below. What follows is by no means a comprehensive catalogue of notable nutrition research contributions from India but only selected examples of contributions which have a direct bearing on major nutritional diseases encountered in the country today.

### MAJOR NUTRITIONAL PROBLEMS

The most outstanding nutritional problems in the country which, today, account for significant impairment of the quality of the country's Human Resources are : Protein energy malnutrition (PEM) leading to physical and mental retardation and underdevelopment of several thousands of children, vitamin A deficiency leading to keratomalacia and nutritional blindness; widespread iron-deficiency anaemia impairing the productivity and increasing the vulnerability of poor populations to infections; and endemic goitre and other iodine-deficiency manifestations affecting physical growth and mental development. With respect to *each* of these problems, nutrition research in India has made important original contributions.

Other nutritional problems of the country which are somewhat more limited in distribution, but which, nevertheless, present fascinating challenges to the biologist and health scientists are : pellagra, in the Deccan plateau; fluorosis in some parts of India, lathyrism in Central India and 'lactose intolerance'. Indian scientific contributions towards the elucidation of the pathogenesis and prevention of these problems also have been no less impressive.

## PROTEIN-ENERGY MALNUTRITION (PEM)

PEM has, as it were, held the central stage in global nutrition research for now nearly three decades. There have been numerous publications on this subject from India, but for our present purpose we may highlight two important Indian contributions which had helped to bring about a major change in prevailing perceptions regarding its pathogenesis and approach to its prevention.

The first was the clear demonstration that the primary dietary deficiency underlying PEM in India was *not* that of protein (as was hitherto being widely claimed) but that of calories.<sup>11</sup> Careful surveys of dietaries of under-fives in different parts of the country among communities in whom PEM was common, showed that while the daily protein intake ranged from 2.8 g/kg body weight to 1.7 g/kg—levels, which on the basis of widely accepted international and national recommendations, could be considered adequate, the daily calorie intakes were of the order of 70 to 75 kcal/kg as against the figure of 100 kcal/kg generally considered adequate. While the dietaries of over 90 per cent of children were deficient in calories, only those of 35 per cent were deficient in protein; even with regard to these latter children, if food intake had been raised to meet their calorie requirement, the protein needs would have been met. There was practically no situation when the children's diets were adequate with regard to calories but deficient with regard to protein alone.<sup>12</sup>

That in the prevailing dietaries of children of poor communities it was the calorie gap that was crucial was further demonstrated in yet another longitudinal study (of 14 months)<sup>13</sup> of a community of poor children whose daily dietaries provided no more than 700 kcal (with 18 g protein daily). In this study it was shown that when the calorie gap in these dietaries was bridged through the supplementation of 300 additional kcal daily, derived from carbohydrate and fat sources (wheat flour, sugar and edible oil) with little additional protein—no more than 3 gms—"empty calories") growth performance could be significantly improved and clinical manifestation of PEM could be averted.

These findings indicated that the prevailing emphasis on "the protein gap" and "protein concentrates" was wholly misplaced; and that the solution to the problem of PEM, fortunately need not have to depend on imports of expensive protein-rich concentrates (fish-protein concentrates were in fact sought to be widely promoted by several

commercial interests), but can be achieved through proper use of inexpensive traditional cereal-legume based diets within the economic reach of poor families and within the country's resources.

It may be pertinent to point out that the foregoing views on the PEM problem, based on research in India, were presented more than 20 years ago at a time when the "protein lobby" was riding high and any suggestion that it was the 'calorie gap' rather than 'protein gap' that was crucial, could be treated as heresay. In retrospect, it will be seen that Indian research contributions, despite the resistance they had initially roused in international circles (Protein-advisory group-PAG) eventually found wide acceptance as evidenced by the great "protein fiasco" and the closure of PAG following thereon. It must be added that Indian scientific contributions which sought to put the PEM problem in proper focus, had found influential scientific support both in developing and developed countries outside India even at the height of the "protein controversy"

### THE PRACTICAL CHALLENGE

The real challenge in the prevention of PEM then boils down to ensuring that children under 5 years of age (especially under 3 years) get their habitual cereal-legume-vegetable foods in amounts adequate to meet their calorie needs—which unfortunately is not the case even at present. The prevailing calorie gap of about 300 k. calories daily in the dietaries of under-fives of poor communities in the country can be bridged with a fraction of the buffer-stocks of foodgrains that we now hold; and yet in the prevailing socio-economic order we are witness to the sad paradox of vast buffer stocks of foodgrains posing storage problems on the one hand, and vast sections of the poor unable to get the food they badly need on the other. Nutrition research has at least helped to expose the stark reality of such inequalities in our present national scene.

Yet another hurdle in feeding cereal-legume based diets to very young children stems from the low calorie-density of these diets—"the bulk factor". Research in India has also attempted to address this issue and to identify traditional home-based techniques through which this bulk factor can be overcome, the viscosity of cooked cereal foods can be reduced and their calorie-density increased.<sup>9</sup>

We may now proceed to consider the second major contribution to the understanding of the PEM problem. It is well recognised that there are two distinct clinical syndromes associated with PEM — namely

'kwashiorkor' and 'marasmus'. (Marasmus in early infancy associated with highly inadequate intakes of milk—'infantile marasmus' has to be considered as a separate category. We are here considering marasmus in the pre-school child). At any given point of time (point-prevalence), it may be computed on the basis of available survey data that while roughly about 1 per cent of under-threes in poor communities may exhibit kwashiorkor, nearly 2 per cent to 3 per cent may show marasmus. Thus, at any point of time, among our poor communities we may expect to see several thousands of poor children suffering from kwashiorkor and several thousands more suffering from marasmus, both of them existing almost side by side in the same villages.

The earlier widely held postulate was that these two manifestations were different diseases with entirely different dietary etiologies—the former due primarily to "protein deficiency and calorie excess" and the latter due to calorie deficiency. If this was really the case we would have had on our hands two major public health problems requiring two entirely different approaches to their prevention and control. Our studies lead us to the conclusion that this is fortunately not the case.

A major contribution of immense practical significance has been the clarification that kwashiorkor and marasmus are *not* two different diseases but just two facets (clinical manifestations) of one and the same central problem of PEM, with a common dietary etiology, and therefore requiring identical approaches for their solution. The differences in the clinical and biochemical features of kwashiorkor and marasmus have been set out in Table I.

On the basis of intensive studies of the actual dietaries of children suffering from these two syndromes and their hormonal profile, we had postulated that 'marasmus' represents the stage of attempted 'adaptation' to the nutritional stress wherein hormonal mechanisms are invoked to ensure that the integrity of highly vulnerable tissues with a high protein-turnover, like the liver, pancreas and viscera, is maintained at the expense of the muscle<sup>11</sup>. Kwashiorkor represents the stage when this 'adaptation' breaks down. Further studies had helped to elucidate the probable nature of the hormonal changes that may be involved in such an 'adaptation' mechanism leading to marasmus at one stage of the disease, and a breakdown of 'adaptation' at a later stage leading to kwashiorkor.<sup>23,24</sup>

**Table I**  
*Kwashiorkor and Marasmus (distinguishing features)*

|  | <i>Marasmus</i>  | <i>Kwashiorkor</i> |
|--|------------------|--------------------|
| Emaciation                             | +++              | +                  |
| Oedema                                 | --               | ++                 |
| Fatty infiltration of liver            | --               | ++                 |
| Serum albumin                          | Almost Normal    | Markedly lowered   |
| Serum Enzymes:                         |                  |                    |
| Lipase                                 | Normal           | Markedly lowered   |
| Amylase                                | Normal           | Lowered            |
| Esterase                               | Slightly lowered | Markedly lowered   |
| Serum Cholesterol                      | Normal           | Lowered            |
| Response to epinephrine                | Exaggerated      | Lowered            |
| Serum urea                             | Normal           | Lowered            |
| Serum Copper                           | Normal           | Lowered            |
| Hair Copper                            | Normal           | Lowered            |
| Urinary urea/total<br>urinary nitrogen | >65%             | <50%               |

Thus in marasmus, elevation of plasma cortisol levels was found to be of a higher order than in kwashiorkor; the adrenal cortical response to injection of corticotrophin was exaggerated. Plasma growth hormone levels and their response to stimuli which were found to be raised in kwashiorkor were not altered in marasmus. Plasma somatomedin activity was found to be low in kwashiorkor but not in marasmus (Table II).

**Table II**  
*Hormonal profile in protein-calorie malnutrition*

|                                  | <i>Marasmus</i> | <i>Kwashiorkor</i> |
|----------------------------------|-----------------|--------------------|
| Plasma Cortesol                  | Very high       | High               |
| Response to ACTH                 | Exaggerated     | Normal or Poor     |
| Plasma insulin                   | Normal          | Normal or low      |
| Response to stimuli              | Normal          | Impaired           |
| Plasma High                      | Normal or low   | High               |
| Plasma Somatomedin               | Normal          | Low                |
| Plasma PBI                       | Normal          | Low                |
| Plasma ADH                       | Normal          | High               |
| Urinary ADH                      | Normal          | High               |
| <i>Source : NIN., Hyd., '74'</i> |                 |                    |

These hormonal changes may help to ensure that in the face of stress posed by nutritional deprivation, muscle tissue is preferentially broken down in order that the structural and functional integrity of more vital tissues like the liver, pancreas and viscera is maintained. Marasmus may thus be looked upon as an extreme stage of 'adaptation'—the farthest limit of what was described as "contraction of the metabolic frontiers". When adaptation eventually breaks down because of continued stress or of its aggravation by super added factors like fresh infections etc., fatty infiltration of the liver, fall in serum albumin, reduction in serum enzymes and oedema ensure with the resultant picture of kwashiorkor. The fact that marasmus and kwashiorkor exist side by side in the same community subsisting on the same diets, as also the fact that marasmus and kwashiorkor could exist *in the same child* at different points of time lend support to the postulate that the two syndromes are but two facets of one and the same disease.

This clarification had helped to place the entire problem of marasmus and kwashiorkor in proper perspective as far as the public health approach to marasmus and kwashiorkor was concerned. We know that we are dealing with two manifestations of a *single* problem and that we do not need two divergent strategies for their control. It is hardly necessary to emphasise here the far-reaching practical implications of this conclusion.

The extensive work done on the foregoing and other aspects of the problem of PEM in India in the fifties and sixties has been reviewed in an earlier publication.<sup>46</sup>

### MISUSE OF THE TERM "ADAPTATION"

In the above discussion we have used the word 'adaptation' to refer to the organism's response to stress. It is, however, important to emphasise that we are *not* using the term adaptation as being synonymous with normalcy and therefore as being something 'acceptable'. Even a severely marasmic child with extreme emaciation but a normal liver is 'adapted'! It is necessary to emphasise this in view of the loose manner in which the term 'adaptation' is now being misused to propagate the view that stunting and 'moderate malnutrition' in Third World children arising from PEM (which are not of such severity as to be life-threatening) may be viewed as acceptable 'adaptation', consistent with their culture ("cultural adaptation") and environment ("Small is healthy"?—For the poor not for



the rich!). It is important in this connection to remind ourselves of the wise note of caution sounded by du Nony as quoted by Kamala S. Jaya Rao.<sup>25</sup> "Adaptation is not progressive, but protective and defensive. When perfect adaptation is attained the organism makes no attempt at further transformation as long as the external stress continues. Therefore ... being adapted .....does not contribute to evolution. Adaptation is never a goal but only a means, a means for *survival*."

It is perhaps not without significance that in keeping with the (implicit) suggestion that "moderate undernutrition" which is not life-threatening may be acceptable 'adaptation', the earlier goal of "child health" is being gradually replaced by the slogan of "child-survival". In several recent publications<sup>15,16,17,18</sup>, the author has cautioned against the danger of the misuse of the concept of adaptation in a manner likely to promote social and political indifference to (and acquiescence in) "moderate malnutrition" in children.

*Keratomalacia and Nutritional Blindness* : The current global approach for the prevention of keratomalacia arising primarily from vitamin A deficiency, through the distribution of two massive annual oral doses of synthetic vitamin A—one each at six-monthly intervals—to children under 3 years of age, was developed and pioneered in India, on the basis of experimental, clinical and field studies. That vitamin A can be stored in the liver for prolonged periods and gets released gradually to meet tissue needs had long been well-established. What was important, however, was (a) to establish the optimal dosage of vitamin A which while not being toxic will be adequate to afford protection to children against keratomalacia for fairly long durations; (b) to identify the most effective and feasible route and form of administration of the vitamin and (c) to demonstrate that under real-life conditions in the field, the administration of vitamin A in the dosage, frequency and form identified as above does help to raise and maintain serum vitamin A levels in children over several months and thus does in fact offer protection against keratomalacia and finally (d) to develop a practical procedure for the routine evaluation of the programme by the public-health agency. Through intensive work carried out in the clinic, laboratories and the field initiated by the National Institute of Nutrition nearly 25 years ago, it became possible for Indian scientists not only to develop this prophylaxis programme but to persuade the planners and policy-makers to incorporate this programme as an integral part of Primary Health Care operations in the country, in the Fourth Five Year Plan itself. The programme is now being continued,

though it must be confessed that its implementation in some States of the country has not been effective—a reflection of the current inadequacies in our Health System. Other countries in the region such as Bangladesh, Nepal and Indonesia are also currently implementing a programme on more or less the same lines.

An indication of the extensive amount of work that was involved in the development of this prophylaxis programme can be obtained from the following brief account of the several steps of the study that preceded the introduction of this programme in the National Health System.

*Choice of Preparation* : In a preliminary trial in which a single dose of 300,000 IU of vitamin A given orally as a water-immiscible preparation to a group of pre-school children, 25 per cent developed signs of acute though transient vitamin A toxicity characterised by raised intra-cranial tension (bulging fontanelles), restlessness, and fever.<sup>47</sup> When the same amount was given as an oil-soluble preparation, the incidence of toxic signs was 4 per cent. Moreover, animal studies had earlier shown that the best hepatic storage was achieved with oral administration of oil-soluble vitamin A.<sup>39</sup> This decided the choice in favour of an oil-soluble preparation.

*Route of Administration* : Oral administration of 10000 IU of oily vitamin A produced significant increase in serum vitamin A levels, but the same dose given intramuscularly had no such effect, as much of the vitamin continued to remain at the site of infection.<sup>39</sup> Oral dosage was therefore preferred as also being more convenient to administer.

*Optimal Dosage Level for Sustained Effect*: Longitudinal studies on groups of children showed that a single oral dose of 300000 IU was able to sustain normal levels of serum vitamin A in children for a period of 6 months. Studies in which 200000 IU of vitamin A was given along with labelled retinyl acetate, and urinary excretion of the label monitored, indicated that 70 per cent of the dose was absorbed and somewhat less than 50 per cent of the total dose was retained.<sup>40</sup> That the administration of such a massive dose was not associated with significant lysosomal damage was demonstrated by the finding that there was either no increase or that an insignificant transient increase in the urinary excretion of lysosomal enzymes arylsulphatase and acid phosphatase following on the administration.<sup>41</sup>

*Field Trial* : The real acid test of the efficacy of this prophylaxis approach consisted in the demonstration through a prolonged field trial extending to 5 years, including 2500 under-fives drawn from several villages, that with 300000 IU of vitamin A administered orally once a year and followed up for a period of 5 years, there was (a) a 75 per cent reduction in the overall incidence of vitamin A deficiency in the community and (b) that there was not a single new case of keratomalacia during this entire period and (c) serum vitamin A levels in children who received the dose were consistently higher than those who had not.<sup>48</sup>

It was after all these extensive time-consuming tests had been completed that the scientists ventured to advise the Government of India to include this programme as part of routine Primary Health Care in at least 9 States of the Indian Union where there was evidence that vitamin A deficiency was more widely prevalent.

As a measure of abundant caution, in order to reduce risk of toxicity to the absolute minimum it was also recommended that the dose be reduced to 200,000 IU at a time and that it be given twice in the year—at six monthly intervals. A practical simple method, feasible under field conditions for the evaluation of the implementation of the programme was also developed.<sup>49</sup>

All this may seem a success story. However, looking back on these efforts which were initiated a quarter of a century ago, and now looking at the results, we may legitimately ask whether all the expectations which prompted these efforts on the part of Indian nutrition scientists have in fact been fulfilled.

The control of nutritional blindness through the 'short-cut' of administration of synthetic vitamin A had been envisaged as a short-term approach—not as the permanent solution of the problem. It was always recognised that the ultimate solution lay in the promotion of the optimal use of B-carotene rich foods—green leafy vegetables—in the dietaries of poor children. It must be confessed that the euphoria and complacency created by the introduction of the prophylaxis through massive dosage of synthetic vitamin A has, to a considerable extent, retarded research designed to develop and promote the better use of inexpensive B-carotene rich foods in the country. If such research has not altogether come to a standstill, it is proceeding, at best, at snail pace as an effort of low priority.

Secondly, the implementation of the prophylaxis programme is obviously tardy especially in States like Bihar. What is most disconcerting is that we do not have any authentic indication as to what real impact the prophylaxis programme has had on the nutritional blindness problem. The official figures of annual incidence cases of nutritional blindness will not stand scientific scrutiny. We do not even seem to have reliable data on changes in the annual incidence of keratomalacia in our leading ophthalmic and paediatric hospitals since the introduction of the programme. In the absence of such data, we are in no position to counter or confirm the claims that are frequently made.

It would seem that while on the one hand we are relying heavily on synthetic vitamin A administration as the answer to vitamin A deficiency which, as pointed out above, was not what was originally intended, on the other hand, we continue to be entirely dependent on a foreign commercial source for our supply of vitamin A concentrate. In short, we are veritably 'at the mercy' of foreign commercial interests with respect to a programme of such vital importance to national health. The claims that a major part of the vitamin A we need for our programmes is being 'indigenously manufactured' will again not stand scrutiny—unless we accept that 'bottling' and participation in the subsidiary and final stages of manufacture amounts to 'indigenous manufacture'; the truth appears to be that the technical knowhow for the *essential step* in the manufacture of synthetic vitamin A rests with a foreign commercial source which virtually holds the monopoly in this regard. It is not clear as to why Indian scientists have shown no enthusiasm in achieving self-reliance in this regard and why no "Technology Mission" has been set up for this valid purpose.

While we may rejoice at our modest successes, we have still a long way to go. Recent research indicates that the implication of vitamin A deficiency may be more far-reaching than what we had earlier imagined. This is all the more reason that this subject should now receive renewed attention.

*Iron Deficiency Anaemia* : A major contribution of immense practical value has been the development of a technique for the fortification of common salt with iron. The research that was involved was not just a simple exercise in food technology but included studies of bioavailability and field trials to determine acceptability and efficacy.

Contrary to the general belief that iron-deficiency anaemia is mostly a disease of women in the reproductive age group, studies carried out

under the auspices of the Indian Council of Medical Research showed that it is also very much a disease of pre-school children and indeed even of adult men. A more recent study by the National Institute of Nutrition<sup>36</sup> showed that 65 per cent of adult women, 75 per cent of pregnant women, 77 per cent of pre-school children and nearly 45 per cent of adult men in poor rural communities were anaemic. Anaemia is thus probably the most extensive nutritional deficiency disorder in the country. Recent research has indicated that apart from impairing productivity, the disease also carries quite a few other functional implications. Though Indian dietaries generally provide 20 to 30mg of iron daily, in view of their high phytate content because of the predominance of cereals, the bioavailability of dietary iron as determined by radioisotope technique is only 1 to 5 per cent.

The rational ultimate answer to the problem would, of course, consist in the diversification and improvement of the diets—a goal unlikely to be achieved in the near future. The programme of distribution of iron-folate tablets through the health system can reach only a small proportion of the population. Under the circumstances, a sensible practical approach would be to increase intake of iron through fortification of a suitable dietary item with iron. Since common salt is a food commodity in universal use and since the poor take it in almost the same amounts as the rich, common salt was the obvious suitable candidate for iron-fortification.

*The Formula* : The real challenge here was to identify the formula for fortification which will satisfy the conflicting requirements of stability, acceptability and bioavailability. The formula which was identified as one satisfying these requirements by the National Institute of Nutrition consisted of *ferro ortho* phosphate (3.5g per kg) and sodium acid sulphate (5g per kg) as an absorption promoter, providing 1 mg iron per g of salt. Later this formula was further improved by the substitution of ferric phosphate by the much less expensive ferrous sulphate (3500 ppm) and orthophosphoric acid or sodium orthophosphate (2800ppm). With an estimated intake of 15gm of common salt per adult per day, common salt fortified as above will provide an additional 15mg of iron.

*Field Studies* : The acceptability and efficacy of salt fortified as above was investigated through a field trial lasting for 18 months among 1600 (boys and girls) school children between 5 and 15 years who were divided into two matched groups, one receiving fortified salt and the other unfortified salt. The culinary acceptability and physiological efficacy of

the fortification procedure were clearly demonstrated<sup>36</sup>. Following on this a multicentric study coordinated by the National Institute of Nutrition, covering a population of 6000 was also undertaken. In this study the salt was made available to the population through the regular food distribution system. Analysis of data on haemoglobin levels in the experimental and control group again helped to confirm the significant impact of the procedure on the anaemia problem (Report of the Working Group 1982). The Government of India have now been persuaded to undertake this programme at least in some parts of the country in the first instance.

An important hurdle that had to be crossed was to find if fortification of common salt with iron was compatible with the decision of the Government of India to resort to universal iodation of common salt meant for human consumption in the country as a method of prevention and control of endemic goitre. Scientists of NIN have now been able to recently develop a feasible procedure for the simultaneous fortification of common salt with iron and iodine.

The above studies may perhaps lack the glamour of some ongoing research in 'advanced' areas of molecular biology and genetic engineering; but their merit consists in the fact that skills in the fields of organic chemistry, biochemistry, physiology and epidemiology were effectively combined and coordinated towards providing a practical solution to major health problems of the country.

*Endemic Goitre and Iodine-Deficiency*: According to some estimates more than 40 million people in the country suffer from goitre. The National Goitre Control Programme based on iodation of common salt which had been initiated in the later half of the fifties, after an initial promising start, had languished because of poor implementation and inept supervision.<sup>19</sup> The emergence of new goitre-endemic areas has added fresh dimensions to the problem.<sup>2</sup>

Recent studies from India have provided important indications of hitherto unsuspected serious dimensions of the problem of neonatal chemical hypothyroidism (NCH) in endemic goitre zones.<sup>26,27</sup> As high as 13 per cent of neonates in endemic goitre areas have been shown to be functionally decompensated on the basis of T4 and TSH levels in their cord blood as determined by the radio-immunoassays techniques. This observation corresponds closely to the finding of a study under the auspices of the Nutrition Foundation of India that nearly 15 per cent of school children investigated in endemic goitre districts showed evidence

of varying degrees of mental under development. These findings have lent urgency and importance to our National Goitre Control programme which has yet to achieve its full stride. A somewhat complacent view of the role of iodine-deficiency in mental underdevelopment had earlier been taken in view of the very low incidence of cretinism and deaf-mutism in the endemic goitre zone.

Parenteral administration of iodised oil to pregnant woman is now being promoted in some quarters as a suitable prophylactic approach in relatively inaccessible areas till such time as the salt iodation programme gathers full momentum. Recent Indian studies<sup>28</sup> however, sound, a note of caution against resorting to this approach. According to these studies, iodised oil injections, when given to mothers particularly in the last trimester of pregnancy do *not* help to reduce the incidence of neonatal chemical hypothyroidism, the "relevance or even the safety" of administering iodised oils to pregnant mothers has been seriously questioned. These views have been challenged and there is apparently some controversy. It must, however, be clear that it will not be prudent to push ahead with any procedure regarding the safety of which serious doubts have been expressed, especially when a time-tested safe inexpensive alternative (salt iodation) is already available.

It is at least gratifying that thanks to our scientists, major decisions on technologies to be opted for in our national public health programmes are not entirely dependent on advice and recommendations of foreign and international agencies.

The foregoing account deals with major nutritional deficiency disorders affecting vast numbers of the country's population. What follows is a brief account of scientific contributions from India towards the better understanding of four other nutritional disorders which while, though not as extensive as those described earlier, are of considerable interest to health and nutrition scientists all over the world.

*Pellagra* : Pellagra is a classical nutritional deficiency disorder traditionally associated with poor populations whose staple is maize (corn). The low content in maize of the essential amino-acid tryptophan, the precursor of nicotinic acid, has been generally held responsible. The important finding from India which ran clearly counter to this well-accepted view was that endemic pellagra in the Deccan plateau of India occurred in populations subsisting not on maize but on the millet sorghum (jowar), which is *not* poor in tryptophan. A feature common to both maize

and sorghum, however, is the high content of the amino acid leucine. This finding sparked off a new series of studies on pellagra starting with a paper<sup>20</sup> in which we had proposed that the high level of leucine in sorghum may play a positive role in the pathogenesis of the disease. The studies which followed showed that excess leucine in otherwise poor dietaries, could induce disturbances in the tryptophan—niacin pathway reflected in increased urinary excretion of quinolinic acid on leucine feeding<sup>5</sup>, decreased rate of synthesis of nicotinamide nucleotides by erythrocytes<sup>37</sup>, decreased activity of quinotinate phosphoribosyl transferase (QPRT) a key enzyme in NAD synthesis in livers<sup>4</sup>, and a fall in platelet 5-hydroxy tryptamine levels<sup>31</sup>.

From these studies it was concluded that excess leucine in poor sorghum diets could bring about significant changes in a number of key enzymes in the tryptophan—niacin pathway ultimately resulting in decreased nicotinamide nucleotide formation from dietary tryptophan—thus leading to conditioned deficiency of nicotinic acid.

In further studies it was shown that these effects of excess leucine could be countered by pyridoxin. Post-tryptophan load excretion of xanthurenic acid, kynurenic acid and quinolinic acid, which were initially raised in pellagrins, were reduced after pyridoxin treatment.<sup>32</sup>

It would thus appear from the Indian studies that in the pathogenesis of pellagra (which is by no means exclusively confined to maize eaters only but which could also occur in sorghum eaters), apart from tryptophan deficiency (in maize eaters) leucine excess (in sorghum-eaters), and deficiencies of pyridoxin and nicotinic acid (in both maize-eaters and sorghum-eaters) may all play a part.

The above observations on the possible role of leucine in pellagra have been contested and challenged by some scientists from Europe and USA. Some recent reports from England have, however, lent support to the observations from India. Magboul and Bender<sup>33</sup> showed that diets which provide excess leucine brought about "significant reduction in the concentrations of nicotinamide nucleotides in liver and blood." The effect was only apparent when the diets provided less than adequate amounts of nicotinamide. The addition of leucine was also shown to bring about "significant activation of tryptophan oxygenase and inhibition of kynureninase." In a subsequent communication Bender<sup>6</sup> reported that "dietary excess of leucine led to inhibition of kynureninase and increased the activity of piconilate carboxylase—which could be expected to explain



decreased synthesis of nicotinamide nucleotides. The very fact that the Indian claims which started with a paper as early as 1960 and which, at one stage, were sought to be dismissed, still continue to attract attention, controversy and support is perhaps, by itself, of some significance.

*Lathyrism* : Neurolathyrism characterised by spastic paraplegia affecting the lower extremities is an ancient disease and is endemic in areas in which diets are predominantly based on the pulse *Lathyrus sativus*. Though the association of lathyrism with the consumption of the pulse has been known for over a century, the toxic factor in the pulse responsible for the disease could not be identified, mainly because the disease could not be reproduced in experimental animals.

A major breakthrough was achieved at the National Institute of Nutrition<sup>43</sup> when it was demonstrated that alcoholic extracts of *Lathyrus sativus* could produce neurotoxic manifestations when injected into baby chicken. The toxic factor was subsequently isolated and identified as BOAA (B-oxalyl amino alanine).<sup>1,35</sup> A simple household method by which the toxin can be completely removed from the seed by steeping the seeds in hot water for about an hour, or by parboiling the seed in a process similar to the parboiling of rice was also developed.<sup>34</sup> Simultaneously attempts were also made by agricultural scientists in India to identify and selectively propagate genetic strains of *Lathyrus sativus* low in BOAA but these attempts have not been successful; but are recently being revived in other parts of the world (Canada).

It must, however, be confessed that all the scientific research efforts that went in into the elucidation of the problem of lathyrism have not directly resulted in the eradication of the disease. Attempts to ban the cultivation of the offending crop failed because the crop had the merit of being a hardy one that would grow on unirrigated land; it was the staple of the poor and there was no easy substitute. Recently, however, it would appear that following on the relative decline in production of pulses in the wake of the green revolution, *Lathyrus sativus* has found a flourishing market as an adulterant of other more expensive pulses like Bengal gram and reportedly is being widely exported out of the endemic zone for this purpose.<sup>21</sup> To the extent to which these new developments dictated by commercial considerations reduce sole reliance by the poor of the endemic regions on *Lathyrus sativus* as their staple food, it might have still done some good, but if the profitability of adulteration should act as an

incentive for intensive cultivation of *Lathyrus sativus*, the problem would be disseminated well beyond the present 'endemic' zones.

*Fluorosis*: While in other parts of the world, there are active movements for fluoridation of water as a method for prevention of dental caries, in India, the problem in some parts of the country (especially Punjab and Andhra Pradesh) is the presence of excess fluoride in drinking water leading to skeletal changes which may sometimes be so severe as to be incapacitating.

Endemic fluorosis was in fact first identified in the country in some areas of the present Andhra Pradesh which were then parts of the erstwhile Madras Presidency.<sup>44</sup> Subsequently, endemic fluorosis belts were also identified in Punjab.<sup>45</sup> The disease affects the rural poor in areas where the drinking water may contain as high as 15 ppm of fluoride. While attempts to defluoridate water using inexpensive adsorbents like paddy-husk carbon have failed to make any significant dent on the problem, recent studies in India have shown that the disease has acquired new serious dimensions. In parts of Andhra Pradesh where the disease has been known to be endemic, it was noticed that large numbers of adolescents and young adults stated developing serious bone deformities generally characterised by marked genu valgum<sup>29,30</sup> manifestations which had never been seen in those areas in earlier years. The prevalence of these deformities ranged from about 2 per cent in some areas to as high as 17 per cent in others and was found to be higher in jowar eaters than those not subsisting on jowar.

A series of interesting studies revealed that this new aggravation of this ancient disease was related to the construction of the large Nagarjunasagar Dam which had impounded large amounts of water. The sequence of events leading up to these new manifestations was as follows: construction of dam and impounding of water—elevation of subsoil water in wide areas in the vicinity of the dam—soil alkalinity—changes in the concentration of trace elements in food-grains in the area and in particular increase in concentration of molybdenum in the foods grown—increased urinary excretion of copper—osteoporosis (super added to fluorosis)—genu valgum. Positive evidence in favour of the actual occurrence of several steps briefly mentioned above has been forthcoming from the several studies at the National Institute of Nutrition.

As a preventive measure, it was suggested that the rural poor should be advised against drawing water for drinking purposes from the wells in the area—the water there being high in fluoride concentration. Instead the

Government was advised that part of the impounded water which was being diverted almost entirely for irrigation purposes through canals should be made available for drinking purposes.

Here is an instance of an unexpected ecological repercussion of a 'developmental programme' which was envisaged as an unmixed blessing that would help to irrigate vast tracts of land and help grow more food.

*Lactose Intolerance* : Chronic diarrhoea arising as a result of intolerance to disaccharides due to deficiency of disaccharidases is now being reported from some parts of the world. The incidence of lactose intolerance is reported to be high among Asians and Africans and rare among Caucasians.<sup>3,7,8,10</sup> On the basis of these findings it was being postulated that inclusion of milk in the diets of undernourished populations of developing countries may lead to undesirable sequelae such as abdominal discomfort and diarrhoea.

Indian studies<sup>38</sup> showed that there was no correlation between signs of 'lactose intolerance' as determined by lactose overloading tests and the levels of the enzyme. It was pointed out that lactose intolerance demonstrated under the artificial conditions of the tolerance tests did not necessarily imply milk intolerance. It was thus shown that there was no case for withholding milk from undernourished Asian children, nor for providing them lactase tablets every time they had a milk drink as was being suggested by some commercial interests. These observations helped to dispel doubts about a traditionally highly valued item of Indian dietaries.

I am deeply aware that the foregoing account does not do full justice to *all* the important recent Indian work in the field of Human Nutrition. For example, we have not discussed the basic contributions of the scientists of the Department of Biochemistry of the Indian Institute of Science and of the National Institute of Nutrition in the field of biochemistry of important vitamins; nor have we discussed the practical contributions of scientists of the Central Food Technological Research Institute (CFTRI) in the field of food technology, particularly with respect to the development of safe, feasible, though not glamorous, methods of food storage technology devoid of potential health hazards. The extensive work in India on human lactation, output and composition of breast-milk; on aflatoxicosis including the demonstration for the first time of the carcinogenic effect of aflatoxin in *primates* and the work on other food toxins; and the more recent work on the amount and nature of 'invisible'

fat in cereals and pulses by Ghafoorunissa and Achaya which has provided new insights with respect to fat requirements in Indian dietaries, have not been mentioned. Perhaps the most important omission relates to what may be the least spectacular but probably the most useful continuing investigations and compilation of the nutritive value of Indian foods—the work which provides the basic data on which all our dietary recommendations rest. However, as I had pointed out even at the start, what has been attempted here is not a presentation of a comprehensive catalogue but of a few selected examples of Indian scientific contributions to the amelioration of undernutrition among our people.

*Concluding Comments* : A complaint that is often voiced is that Indian scientists generally tend to pursue research on problems which are in current fashion in technologically advanced countries rather than on those which are of immediate relevance and practical importance to their own country. These latter problems may seem to lack 'glamour' and 'visibility' and may not offer promise of "international recognition". They may not also call for ultra-sophisticated techniques—the use of which by itself is often mistakenly assumed to be a hallmark of distinction. Under the circumstances, it should not be surprising if the anxiety 'keep in step' with leaders "elsewhere" and to be considered as being in the mainstream of currently "hot" international scientific pursuits, often takes precedence over obvious national needs.

The foregoing account of Indian nutrition research contributions over the last few decades, however, will show that these criticisms will not apply at least to Indian scientists engaged in research in the field of Human Nutrition. Nutrition scientists have, by and large, tried to address the major problems right at their doorstep and the entire thrust of their research had been directed towards understanding the evolution of these problems and more importantly towards identifying practical solutions to them—solutions which can be applied given the present constraints. It must also be remembered that the overall budgetary allocation to nutrition research today represents a tiny fraction of the allocation to several other sectors which command highly articulate advocacy and visibility.

With respect to many areas of scientific activity there is a gap between the accumulation of knowledge in research laboratories and the practical application of that knowledge in the field for the good of the people. In no field of scientific activity is this gap perhaps greater (and more unfortunate and inexcusable) than in the field of Human Nutrition.

Today we have the necessary technical knowledge gathered through painstaking research—and, to a great extent, also the resources—with which we can eliminate at least our major nutritional deficiency orders—even if we are in no position to ensure for our people a level of nutritional status comparable to those of advanced prosperous countries. But the implementation of practically every major nutrition programme—be it control of PEM, goitre, nutritional blindness, anaemia, lathyrism or fluorosis—is tardy and inefficient. A country which can join the race for exploration of space and expeditions to the Antarctic, which boasts of very impressive 'scientific manpower' and is "among the top ten" industrial powers of the world, apparently finds itself unable to summon necessary political will and administrative competence to wipe out diseases which needlessly afflict millions of her people—diseases which have long ago been eliminated from the rest of the civilised world and for the elimination of which the scientific know-how and requisite material and manpower resources are available within the country.

It is not that we scientists have all the answers; nor is it that our answers are all necessarily sound and feasible. Our accomplishments do not merit wild jubilation but justify some optimism. There is undoubtedly considerable scope and need for continued maintenance, and indeed further improvement, of the quality and range of nutrition research in the country. There is absolutely no room for complacency; the unfinished tasks that lie ahead are truly formidable. While practically all the 'old' problems of undernutrition are still very much with us, 'new' problems are emerging. The 'successful' green revolution has unleashed a major deterring repercussion—namely the progressive reduction in per capita availability of pulses (legumes), the poor man's protein. There is reason to believe that for this reason, the *quality* of the dietaries in our poorest households has actually deteriorated sharply during recent years in the wake of intensive agricultural technology, ill-monitored changes in soil chemistry have been induced and these may be expected to be reflected in distortions of trace element composition of foods; new goitre-endemic areas are being identified; with the demographic transition nutritional problems of the aged are coming to the fore; with increasing urbanisation, nutritional problems of urban slums are gathering in intensity. Since our current health programmes largely stop with "death-control strategies" ('child survival' and 'oral rehydration'), we may expect that marginal reductions in child mortality will be accompanied by increasing prevalence of childhood malnutrition in the next two decades. At the other

end of the socio-economic spectrum we may also expect increasing incidence of degenerative diseases brought on by increasing life expectancy and overnutrition associated with affluence. All this is not to paint a pessimistic picture of the future but merely to forewarn that the challenges which the nutrition scientists may face in the next two decades may prove even more formidable and exacting than those of the last three decades. For this reason, Nutrition Research must receive much higher priority and support in our national scientific agenda in the future than in the past.

## References

- 1 Adiga P R, Rao S L N and Sharma P S 1963 Some structural features and neurotoxic action of a compound from *L sativus* seeds; *Curr Sci* 32 153
- 2 Agarwal K N, Agarwal D K and Srivastava S 1984 Emergence of goitre in Delhi, *NFI Bull* 5 6-7
- 3 Bayless T M and Rosenweig N S 1966 A racial difference in incidence of lactase deficiency. A survey of milk intolerance and lactase deficiency in healthy adult males, *J Amer Med Ass* 197 968-972
- 4 Belavady B and Gopalan C 1965 Production of black tongue in dogs by feeding diets containing jowar (*Sorghum vulgare*); *Lancet* 2 1220-1221
5. Belavady B L, Srikantia S G and Gopalan C 1963 The effect of the Oral Administration of Leucine on the metabolism of Tryptophan; *Biochem J* 87 652-655
6. Bender and David A 1983 Effects of a dietary excess of leucine on the metabolism of tryptophan in the rat: a mechanism for the pellagrigenic action of leucine; *Br J of Nutr* 50 25-32
- 7 Bolin T D and Davis A E 1969 Asian lactose intolerance and its relation to make of lactose; *Nature* 222 382
8. Cook G C and Kajubi S K 1966 Tribal incidence of lactose deficiency in Uganda; *Lancet* 1 725-729
9. Desikachar H S R 1981 Malting as an aid in reduction of viscosity of cereal and legume based diets; *NFI Bull* April, p-6
10. Flatz G, Saengndom C and Sanguabhokhai T 1969 Lactose intolerance in Thailand; *Nature* 221 758
- 11 Gopalan C, Kwashiorkor and Marasmus 1968 Evolution and Distinguishing Features; 49-58 eds R A McCance and E M Widdowson (London : J & A Churchill Ltd.)

- — and Narasinga Rao B S 1971 Nutritional constraints on growth and development in current Indian dietaries; *Proc Nutr Soc India*, **10** 111
- —, Swaminathan M C, Krishnakumari V K, Rao Hanumantha D, Vijayaraghavan K 1973 Effect of calorie supplementation on growth of undernourished children; *Amer J Clin Nutr* **26** 563-566
- — 1983 Small is healthy? For the poor; not for the rich? *NFI Bull*, 1-5
- — 1982 The Nutrition Policy of brinkmanship; *NFI Bull* 5-6
- — 1984 Child Survival and Child Nutrition; *NFI Bull* **5** 1-3
- — 1988 *Stunting: Significance and implications for Public Health Policy in linear growth retardation in less developed countries*; ed John C Waterlow, Nestle Nutrition workshop series Vol 14 266-269. (New York: Nestec Ltd., Vevey/Raven Press Ltd.)
- — 1989 "Undernutrition: Measurement and Implications" Nutrition, Health and National Development *NFI Special Publication Series* **4** 69-100
- — 1981 The National Goitre Control programme A sad story; *NFI Bull* 1-2
- —, Srikantia S G 1960 Leucine and pellagra, *Lancet*, **1** 954-957
- — 1984 Scientific Report - 2, *NFI* 54-55
12. Gopaldas T, Mehta P, John C 1989 Reducing the bulk of Cereal Weaning Gruels - A simple technology for rural homes; *NFI Bull* **9** 5-8
13. Rao K S J, Srikantia S G and Gopalan C 1968 Plasma Cortisol levels in protein-calorie-malnutrition; *Arch Dis Child* **43** 365
- — 1974 Evolution of Kwashiorkor and marasmus; *Lancet* **14** 709-711
- — 1977 Nutritional stress and hormonal adaptation; *Q J Surg Sci* 306-313
14. Kochupillai N, Yalow R S 1978 Preparation, purification and stability of high specific activity <sup>125</sup>I - labeled thyroxines; *Endocrinology* **102** 128-135
- — 1984 Neonatal thyroid status in iodine deficient environment of sub Himalayas region; *Indian J Med Res* **80** 293-299
- — and Godbole N M 1986 Iodised Oil injections in Goitre Prophylaxis: Possible impact on the new-born; *NFI Bull*, **7** 1-3
15. Krishnamachari K A V R and Krishnaswamy K 1974 An epidemiological study of the syndrome of genu valgum among residents of endemic areas for fluorosis in Andhra Pradesh; *Indian J Med Res* **62** 1415
- —, Krishnaswamy K, 1973 Genu Valgum and Osteoporosis in an area of endemic fluorosis; *The Lancet* **2** 877-879
16. Krishnaswamy K, Rao S B, Raghuram T C and Srikantia S G 1976 *Am J Clin Nutr* **29** 177-181

- — 1970 *Role of Pyridoxine in Pellagra* - V N Patwardhan Prize Oration, ICMR bulletin Vol 9 No 1 P1-3
17. Magboul, Bahieldin I, Bender and David A 1983 The effect of a dietary excess of leucine on the synthesis of nicotinamide nucleotides in the rat.; *Br J Nutr* **49** 321
  18. Mohan V S, Nagarajan V and Gopalan C 1966 Simple practical procedure for the removal of toxic factors in *L. Sativae*; *Indian J Med Res* **54** 410
  19. Nagarajan V, Roy D N, Mohan V S and Gopalan C 1963 Annual Report, *Nutri Res Lab* 39-40
  20. Narasinga Rao B S 1981 Control of anaemia by fortification of common salt with iron.; *NFI Bull.*, 7-8  
 — — and Sivakumar B 1972 Studies on Vitamin A absorption; *Indian Pediatr* **9** 307-310  
 — — and Mohanram M 1971 Urinary excretion of lysosomal enzymes in hypovitaminosis and hypervitaminosis A in children; *Int J Vitam Nutr Res*, **41** 321-326  
 — — Raghuramulu M, Srikantia S G, Narasinga Rao B S and Gopalan C Nicotinamide Nucleotides, *Bio Chem J*; **96** 837-839
  21. Reddy V and Pershad J 1972 Lactose intolerance in Indians; *Am J Clin Nutr* **25** 114  
 — — and Srikantia S G 1966 Serum Vitamin A in kwashiorkor; *Am J Clin Nutr* **18** 34-37
  22. Report of the Working Group on Fortification of salt with Iron 1982; *Am J Clin Nutr* **35** 1442-1451
  23. Roy D N, Nagarajan V and Gopalan C 1963 Production of neurolathyrism in chicks by the injection of *Lathyrus sativus* concentrates; *Curr Sci*; **32** 116-118
  24. Short H M, Mc Robert G R, Bernard T W and Nayar A S M 1937 Endemic fluorosis in Madras Presidency; *Indian Med Gaz*, **72** 369-398
  25. Singh A and Jolly S S 1961 Endemic fluorosis: *Q J Med.* **30** 357
  26. Srikantia S G 1969 Protein calorie malnutrition in Indian children; *Indian J Med Res* **578** (supple.) 36-53
  27. Swaminathan M C 1971 Prevention of vitamin A deficiency by administration of massive doses of vitamin A; *Proc. Asian Congr. Nutr* 1st 696-701  
 — —, Susheela T P and Thimmayamma B V S 1970 Field prophylactic trial with a single annual oral massive dose of vitamin A; *Am J Clin Nutr* **23** 119-122
  28. Vijayaraghavan K, Naidu A M, Rao N P and Srikantia S G 1975 A simple method to evaluate the massive dose vitamin A prophylaxis programme in pre-school children; *Am J Clin Nutr* **28** 1189-1193





**Ashesh Prosad Mitra** (b. 21 February 1927) did D. Phil from (1955), University of Calcutta. He is Bhatnagar Fellow (CSIR), National Physical Laboratory, New Delhi, Secretary, Department of Scientific & Industrial Research and Director General, Council of Scientific and Industrial Research; Director, Scientist (Director's grade), Head, Radio Science Division, National Physical Laboratory, New Delhi

Mitra's major work has been in the field of earth's near-space environment, especially in ionosphere, and aeronomy, through ground-based and space techniques, and pioneering research in the emerging areas of ionospheric chemistry, stratosphere-mesosphere coupling, and the crucial role played by minor constituents in the lower ionosphere and mesosphere and in environment control. His pioneering work on cosmic radio noise for studying the upper atmosphere led to a series of discoveries in ionosphere, solar physics and cosmic rays. He has contributed substantially to the establishment and operation of the International Spacewarn System and the International Ursigramme and World Day service. Over the last decade, Mitra has concentrated on the scientific aspects of global environmental hazards of human activities. His contributions to ozone problem have been far-reaching. He has prepared an inventory of greenhouse gases for India.

Mitra is Fellow of Royal Society of London and Indian Academy of Sciences; President, International Union of Radio Science (URSI), Member, International Academy of Astronautics, and American Geophysical Union, Secretary, INSA (1979-82). He is the recipient of Shanti Swarup Bhatnagar Prize (1968); Kariamanikkam Srinivasa Krishnan Memorial Lecture Award (INSA) (1975); Jawaharlal Nehru Fellow (1978- 80), C.V. Raman Award (UGC) (1980); FICCI Award (1982), Om Prakash Bhasin Award (1987); C.V. Raman Medal (INSA) (1991); G.M. Modi Award (Gujar Mal Modi Science Foundation) (1992); Meghnad Saha Award (National Academy of Sciences, India) (1991), Padma Bhushan (1989).

*Ashesh Prosad Mitra was elected to the fellowship of the Academy in 1961.*

**Address:** Bhatnagar Fellow (CSIR), National Physical Laboratory; Hillside Road, New Delhi 110 012, Tel: (011) 5745298 (Off), 6474357, 6474018 (Res).

## RADIO SCIENCE : J C BOSE TO NOW

A P MITRA FNA

The scope and contents of the field "Radio Science" have changed drastically over the years. In the early years, after the successful communication experiment of Marconi over the Atlantic and the discovery of the Ionosphere on both sides of the Atlantic by Appleton in England and Breit and Tuve in the USA in 1925-26, one understood by the words "radio science" only the problems of the ionosphere and the propagation of radio waves through this ionized medium. The second major field of interest in these early years was "Atmospherics", but only as interference to radio communication. Later these were used for solar flare patrol, for detection of atmospheric nuclear explosions and with the dramatic announcement of the "*whistlers*" by Ratcliffe in 50's—a spectacular aspect of atmospherics for magnetospheric studies. The whistlers are a result of dispersion of VLF radio waves produced by lightning discharges travelling along magnetic field lines from one hemisphere to the other.

Tropospheric propagation of radio waves and the role of scattering and tropospheric ducting came to be recognised as early as 1943 with the unexpected behaviour of 100 MHz radars in Malta and soon, at the close of World War II, a dialogue started between radio scientists and meteorologists resulting in the new field of "*Radio Meteorology*". Around the same time came a new discovery of radio waves signals, i.e., the discovery of radio waves emitted by natural processes from astronomical objects. Thus a new branch of radio science—*Radio Astronomy*—was born.

In techniques the most important was the development and theoretical understanding of masers, of solid state semi-conductor devices, of achieving miniaturization of a scale undreamt of earlier, optical devices and optical communication and now biophotonics. A virtual communication revolution has emerged through a combination of satellite links, television, fibre optics and high speed communication, leading to a video communication revolution which can only be compared, in sheer impact, with the publication of Gutenberg Bible. With the advent of satellites another branch of radio science was born : *Remote Sensing Techniques* opened up for sensing-through radio waves sent from ground or received at ground or signals received from a satellite or received at

satellite—the conditions of the land, the ocean, the ice, the atmosphere, the ionosphere and the magnetosphere.

Thus over the years Radio Science has been widening its scope in many directions, now means a wide variety of things : standards, fields and waves, signals and systems, electronics and optical devices and applications, electromagnetic noise and interference, remote sensing, wave propagation through the ionosphere and troposphere, waves and plasma, radio astronomy, bioelectromagnetics. The ICSU body concerned with radio science—the International Union of Radio Science, URSI—has accommodated these different areas, as and when these emerged, through separate commissions. The newest commission to be formed is commission K on "Electromagnetics in Biology and Medicine" (see Fig. 1).

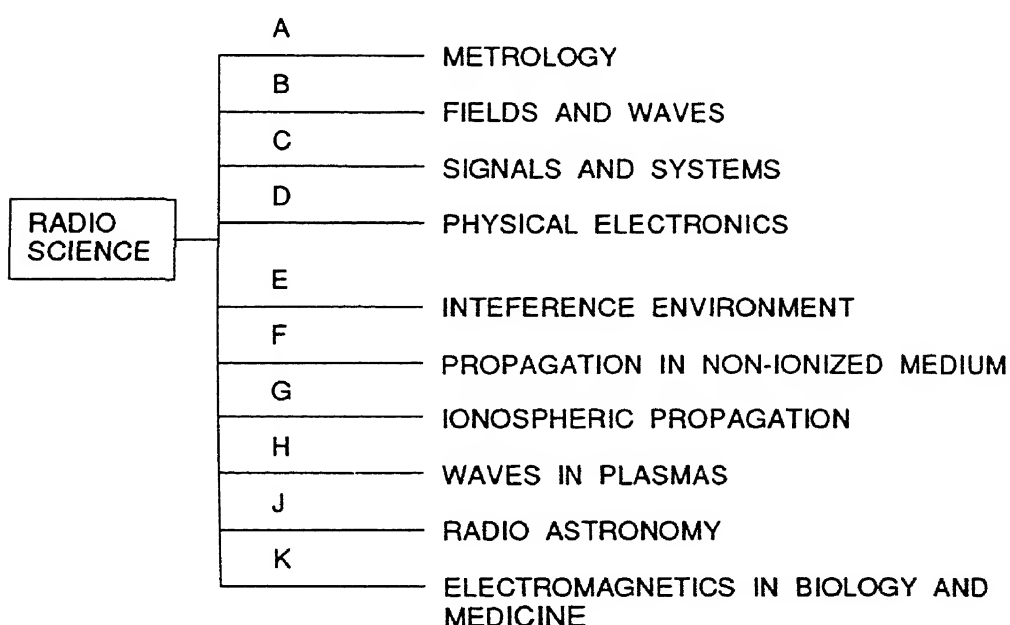


FIG 1. URSI and Radio Science

India has a very distinguished record of work in many of these areas. Over some 100 years there have been singular achievements—all world class. In this presentation, I will take you on a journey of these hundred years of distinguished record of work, focussing on several milestones.

The milestones, relative to major international events, are outlined in Fig. 2.

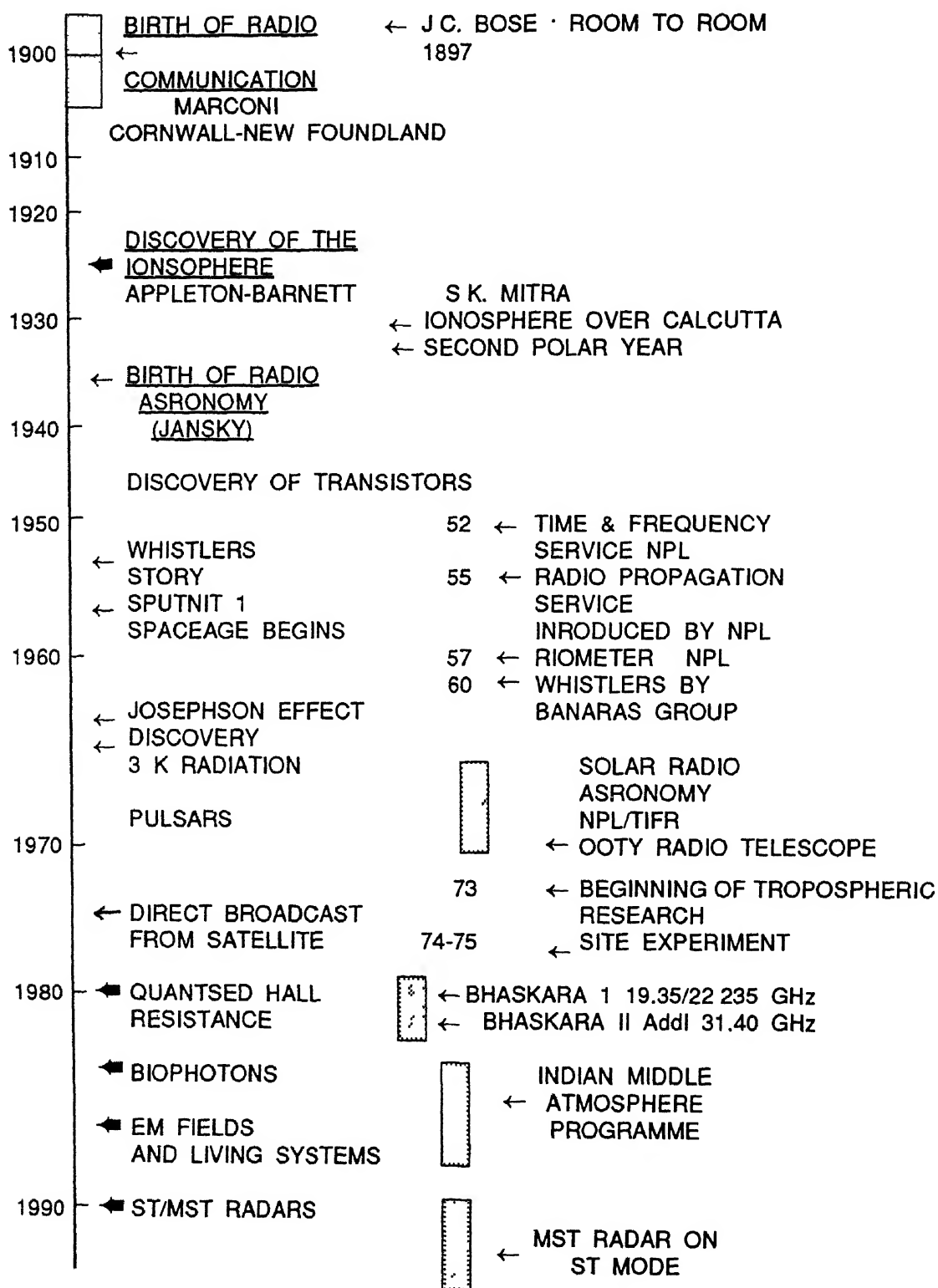


FIG 2. Sequence of major events

## HUNDRED YEARS AGO—THE PIONEERING WORKS OF J C BOSE

The first most significant event in radio science in India, described in several communications in the *Proceedings of the Royal Society* between 1897 and 1902, was a series of classic experiments on generation, reception and optical properties in the then unexplored microwave region of 5cm to 1 cm by J C Bose<sup>24</sup> (Fig. 3).

Wishing to bridge the gap between the long Hertzian waves (a few meters) and long IR waves obtained by purely optical methods (1/20 mm in length), he built a Park Oscillator to produce microwaves, wire gratings to measure their wavelengths and galena crystals as detectors and conducted experiments such as reflection, refraction and transmission through materials such as brick. He also worked on semiconductor materials and on photoconductivity. But perhaps the most important event was his demonstration in 1895 of wireless transmission through solid walls and his repeat experiment at the Royal Institution in London in 1896 a year in advance of Marconi's famous experiment. Soon thereafter his attention moved to plants when in 1900 he discovered, to his surprise, a similarity in the abortive responses of organic and inorganic matter to centimeter wavelength radiators. This, one may say, was the beginning of bio-electromagnetics.

## S K MITRA—THE BEGINNING OF IONOSPHERIC RESEARCH

After that there was a lull for over twenty years, until the establishment of the Wireless Laboratory at Calcutta. Serious work on ionosphere, however, began only around 1930, with a medium wave transmitter, made available by the Calcutta station of the Indian State Broadcasting Service and installing a receiving system some 75km away, providing the first experimental evidence in India of the E-region of the ionosphere. This was achieved only a few years after the discovery of the ionosphere. A series of papers came in quick succession relating to the behaviour of ionospheric layers over Calcutta. It was remarkable that the quality of ionospheric mapping achieved at that time with such simple equipment was so excellent<sup>32</sup>.

This was an exciting period for ionospheric science in India. The timing was also right. The Second Polar Year Programme (IPY2) was being organized, and a study of the Ionosphere was being included for the first time. Mitra (S.K.) decided to formally participate in this programme: this was India's first entry into organized international science. There was

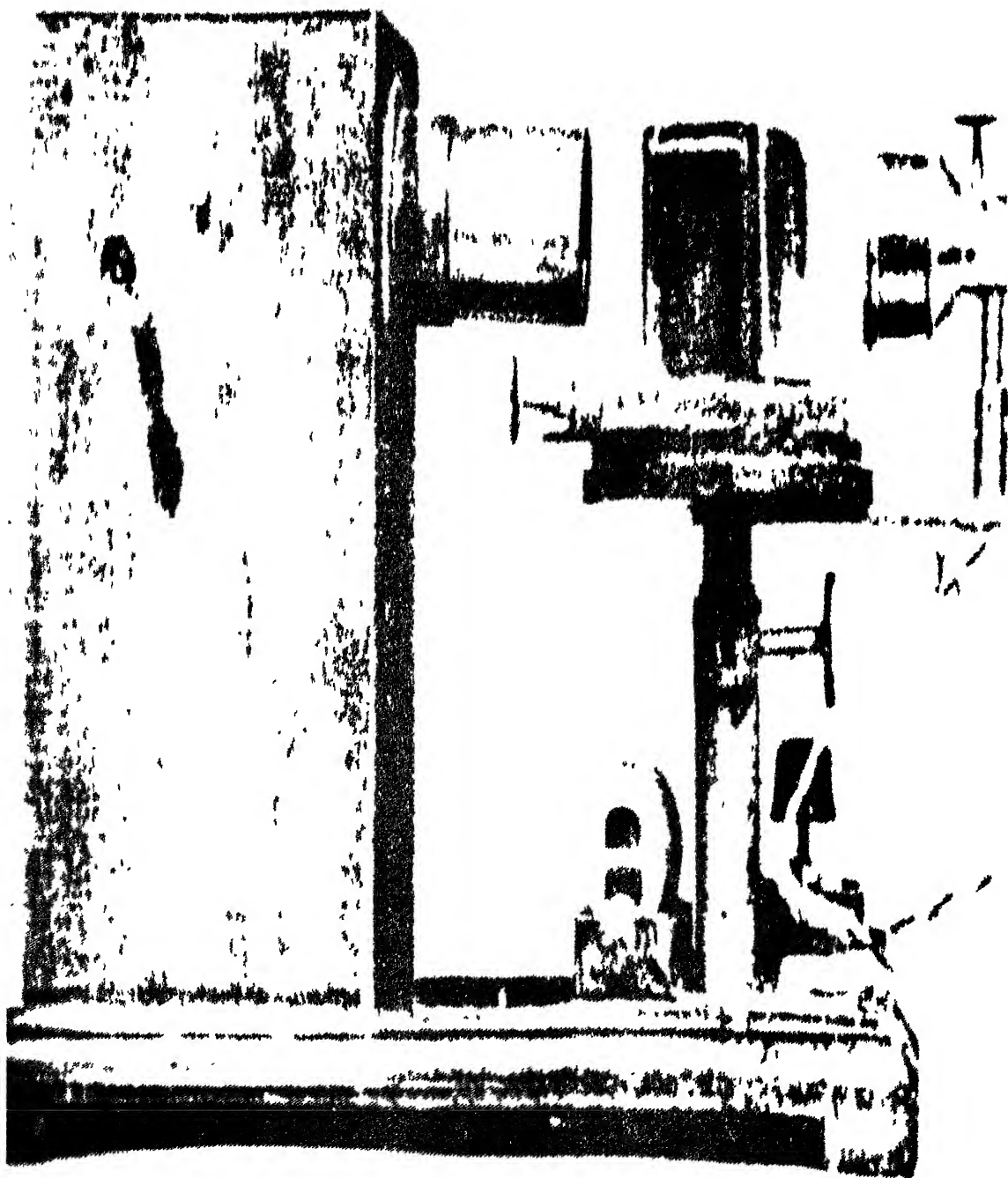


FIG 3 Microwave equipment of J C Bose

no clear idea about the relative contributions (if any) of energetic particles from the sun and the electromagnetic radiations. An excellent opportunity soon came that allowed a distinction between ionization produced by solar electromagnetic radiation and solar corpuscles. This was the occurrence of the annular solar eclipse visible in Calcutta on August 21, 1933. Observations of critical frequencies by varying the transmitter frequency on the eclipse day as well as on the preceding and following days showed a trough of reduced electron density values in the E-region some 20 minutes after maximum observation. Taking into account the natural sluggishness of the ionosphere, this was interpreted as evidence of E-region ionization being primarily (and perhaps exclusively) due to solar UV radiation. This was a correct conclusion, though the timelag of 20 min. was somewhat unusual<sup>1,2</sup>.

Scientists were already looking for existence of additional layers in the Ionosphere. Mitra and Shyam<sup>3</sup> announced in 1935 the detection of regular echoes from low heights (55km); later even echoes as low as 20km (20-30km range) were reported. Reflections from heights around and above 55km were later to be observed extensively with HF transmissions, particularly in Australia, USA and Canada, and the technique came to be known in later years as *Partial Reflection Technique*. The very low level reflections that S K Mitra and his colleagues<sup>30</sup> saw from 20-30km and also reported soon afterwards were, however, not taken seriously for a long time until the concept of using HF as atmospheric radars came up in the sixties. In that sense these early pioneering works can be treated as the forerunner of HF radars.

The manual ionospheric recorder, built and operated by the Calcutta group during that period is shown in Fig. 4.

The *Upper Atmosphere*<sup>36</sup>, which turned out to be one of the best read books in this area came out one solar cycle later in 1947 and its revised edition in 1952.

The *Upper Atmosphere* was totally different from 1935 report<sup>3</sup> (prepared for the then National Institute of Science) although partially based on it. The most important difference was that it considered for the first time the atmospheric environment as a whole—neutral and ionized—its thermal structure and distribution of constituents, its motions, the interaction of the solar radiation and particle stream with these gaseous constituents, the mechanism of airglow. The ionosphere was treated as only a part of this vast panorama that interlinked in a myriad ways the sun,

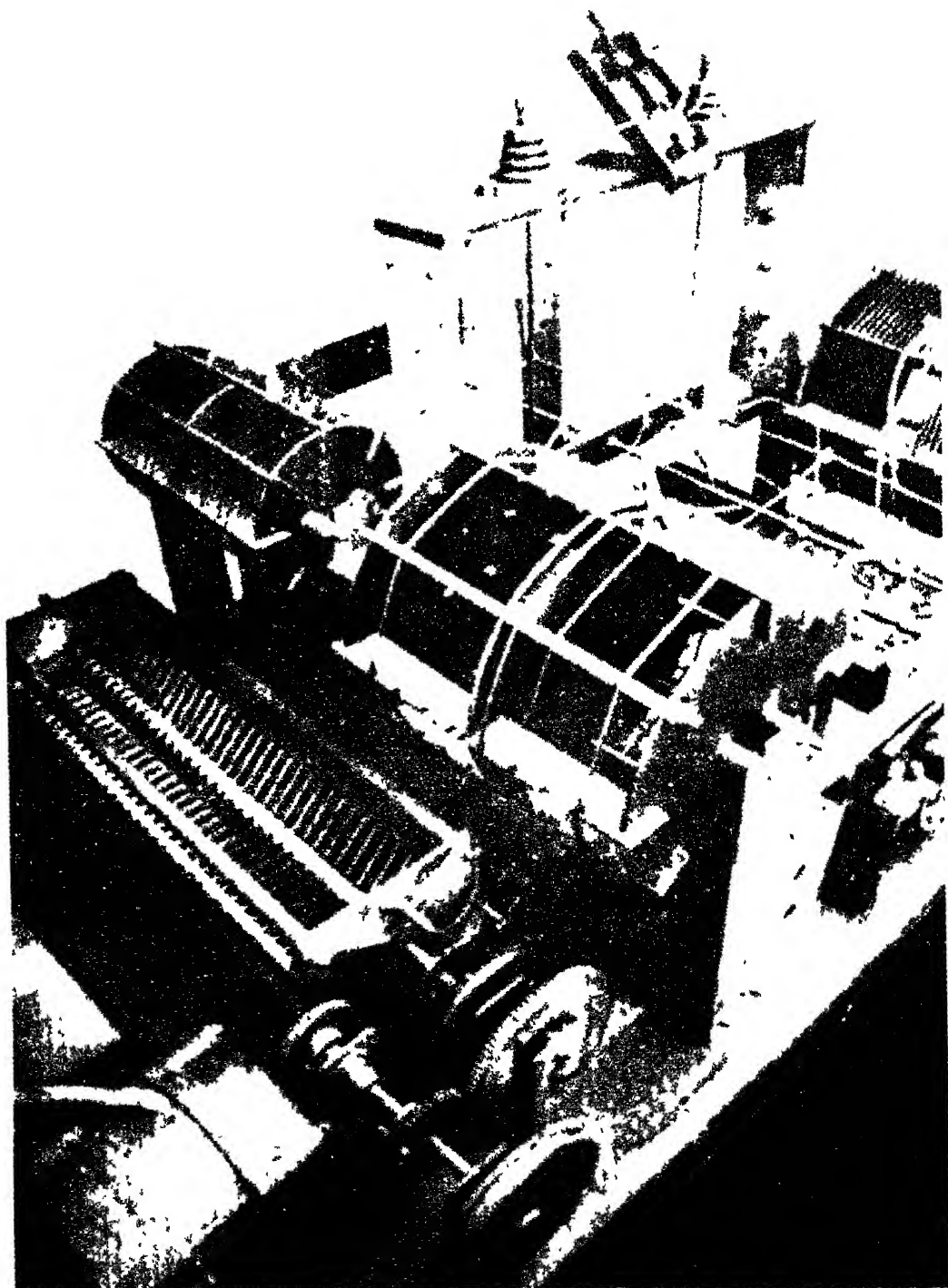


FIG 4. The manual ionospheric recorder built and used by the laboratory of Professor S K Mitra in the forties and fifties



the earth and the atmosphere. This was then an entirely new concept. Secondly, deviating from the then existing practice of studying the ionosphere from the point of view of propagation of radio waves, he viewed the exploring radiowave as a remote sensing tool of the upper atmospheric levels which could not be reached with balloons and were only beginning to be explored by rockets.

### FORTIES AND FIFTIES

J C Bose's pioneering work on microwaves during the turn of the last century lay dormant for some fifty years, until the late forties, when S K Chatterjee and his colleagues<sup>4</sup> at the Indian Institute of Science in Bangalore resurrected microwave research in India. They initiated measurements, for example, on the dielectric properties of materials, artificial sea water and electron transit-time effects in tubes at VHF.

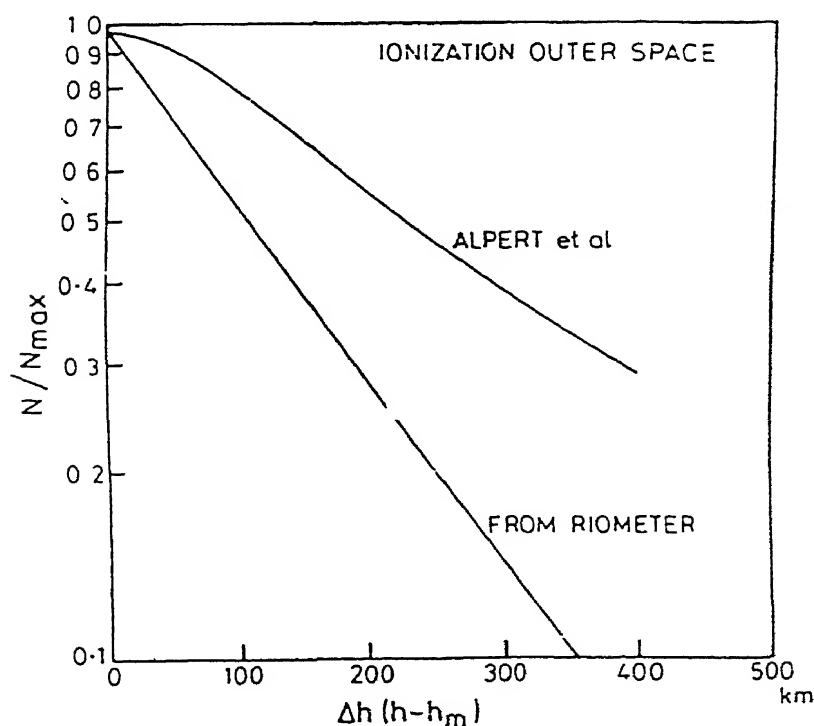


FIG. 5. The first glimpses of the outer ionosphere from riometer and satellite radio signals. The riometer results are from NPL scientists (after Sarada and Mitra)

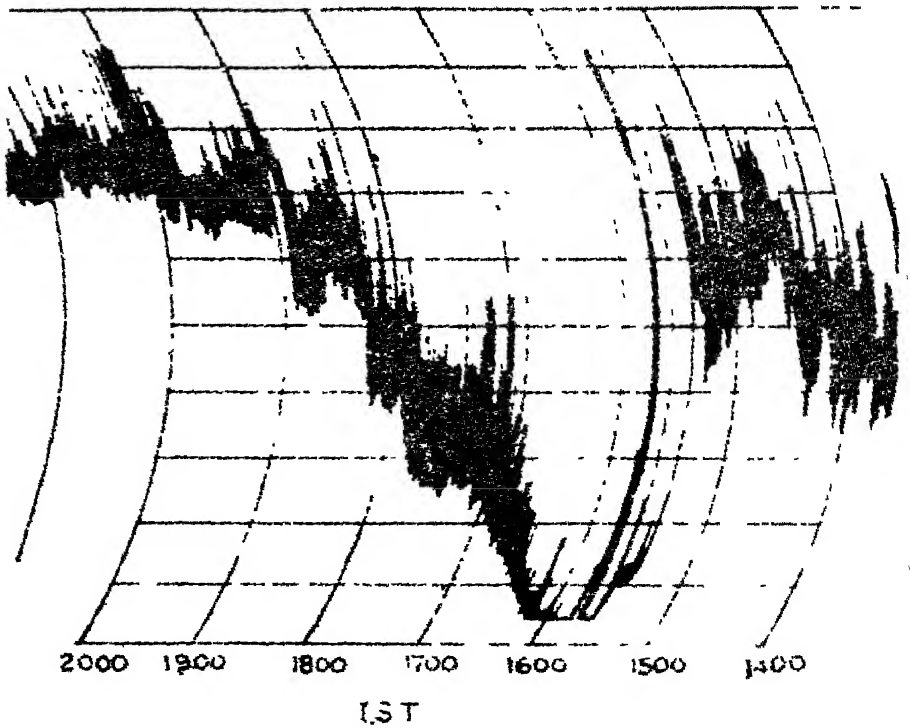
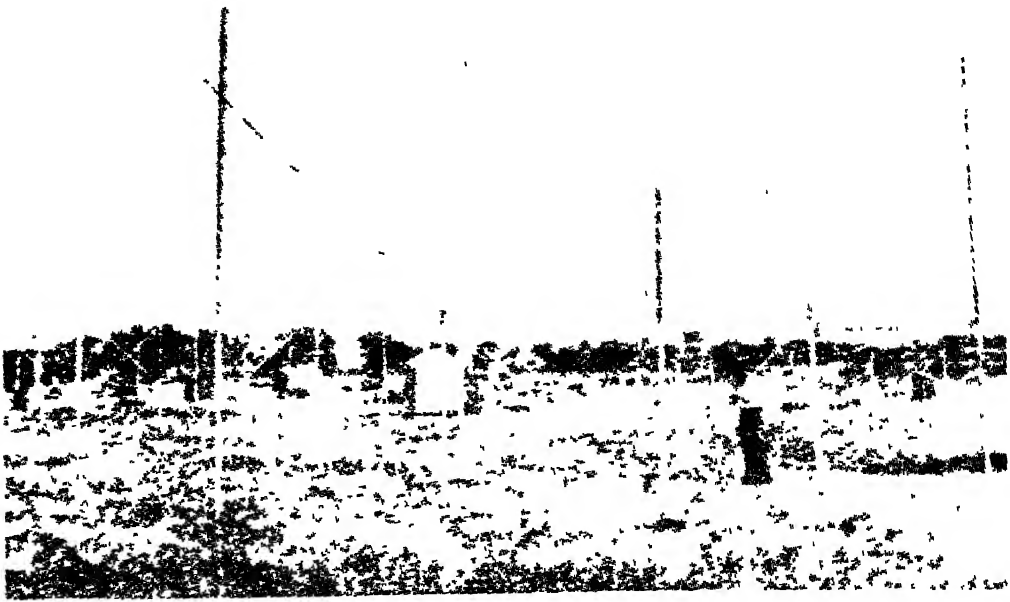


FIG. 6 (a) Original cosmic radio noise equipment (antenna and the hut housing the receiving system shown) installed at the NPL during IGY  
 (b) A severe case of Sudden Cosmic Noise Absorption (SCNA) following a major solar flare recorded with his equipment on March 23, 1958.

The second most important effort was the initiation of developmental work on vacuum tubes upto the pilot plant stage in the laboratory of Professor S K Mitra. This was the beginning of research on electron devices in India, although the initial stages were not very productive<sup>25</sup>. Some of the early products were rectifier tubes—the type 80 diode and type GC5 triode—the first to be made in India, but the enthusiasm was impaired by the Government of India's decision for foreign collaboration for production of electron devices and the discovery of the transistor by Bell Telephone Laboratories shortly afterwards. A major series of work came from K S Krishnan, S C Jain and others on thermionic emissions and, in particular, on measurement of thermionic constants using thermal effusion. Interest soon shifted to high power devices and perhaps the most notable activities were those of Amarjit Singh (1955)<sup>5</sup> on the development of X-band TW magnetrons initially at NPL and later at Pilani in CEERI. The CEERI efforts were later to develop into a major international level expertise for the development of specific high power magnetrons, klystrons and a new generation of microwave devices, gyrotrons, based upon fast-wave interaction, with application in plasma fusion.

### RIOMETER

A major achievement was the introduction of the Riometer. Recognising that use of radio frequencies must in future extend beyond the limited range of 500kHz to 20MHz, a new technique which had been developed a few years earlier by Mitra and Shain<sup>6</sup> in Australia was introduced in India. This was the use of radio noise from our galaxy—the so called cosmic radio noise—at a frequency exceeding but close to the critical frequency of the F-region of the ionosphere. the technique proved to be one of the most versatile ground based techniques for geophysical studies. This was later extensively used during the IGY and became a major tool for the study of the ionosphere and the sun and in particular in high latitudes for detection of Polar Cap Absorption Events (PCAs) caused by low energy cosmic rays impinging on the atmosphere immediately after solar flares. It offered the first glimpse of the outer ionosphere, soon to be a matter of special attention with the launching of Soviet satellite Sputnik-I (Fig. 5).

The Riometer was essentially an Indian experiment. Apart from the Indian contribution in the development of the technique (the early facility established in NPL is shown in Fig. 6(a)), its extensive use in India during and after the IGY for both normal absorption and for flare recording (one

of the largest recorded on the day the IGY began is shown in Fig. 6(b)) was on a high key. While in the high latitudes, the main interest centered eventually on the detection of solar proton events, in India interest continued also in the measurement of normal absorption taking advantage of the very high F-region critical frequencies in these low latitudes. Riometer also turned out to be powerful monitoring tool for atmospheric nuclear explosions and was effectively used by Saha, Karabin and

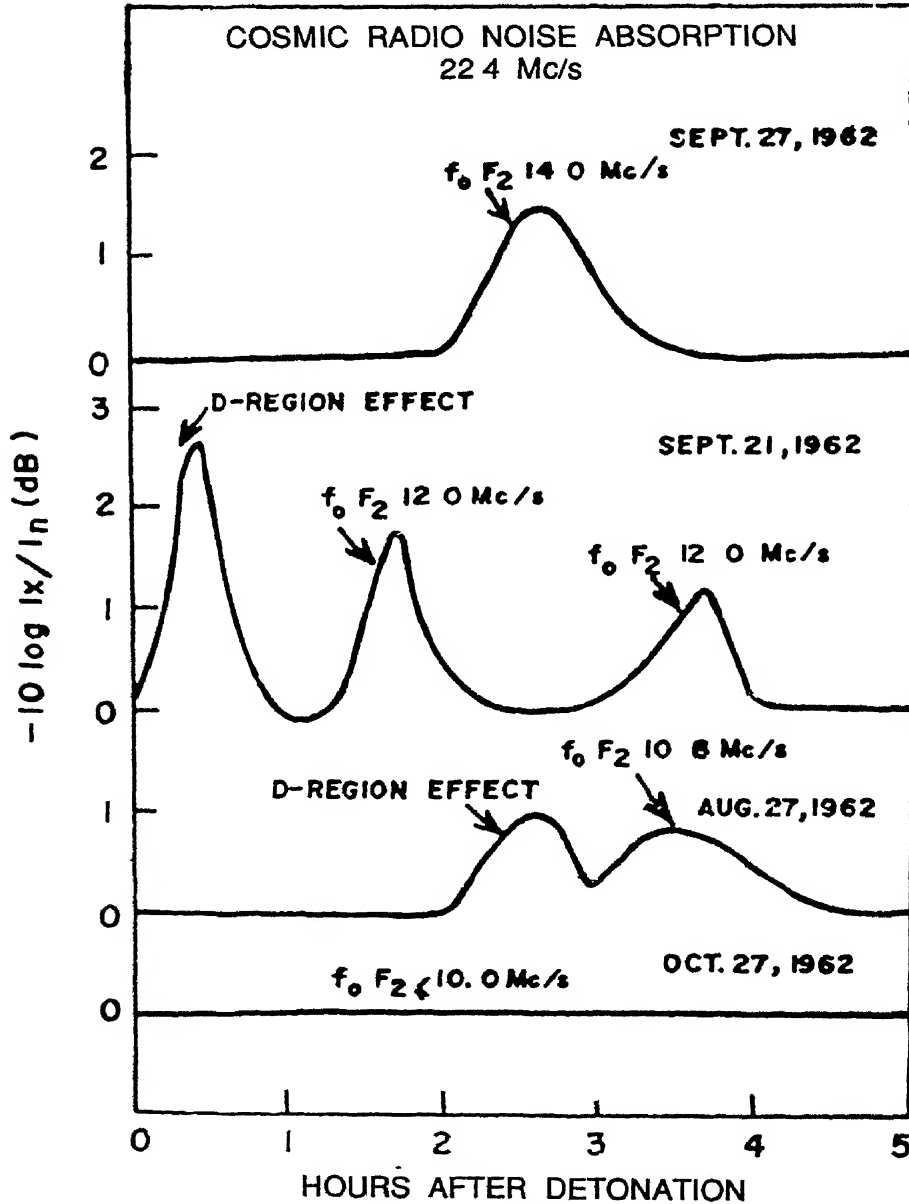


FIG. 7 Detection of atmospheric nuclear explosions with riometer and other ionospheric techniques. Note the separation of D (~80km) and F region (~300km) effects (after Saha & Mahajan, 1964)

Mahajan<sup>7</sup> in the NPL for the U.S. and Soviet tests during 1961-62. A remarkable advantage, from the point of view of understanding the nature and magnitude of upper atmospheric effects of nuclear explosions, was the fact that there was a time separation of the effects in the lower (~80km) and the upper atmosphere (~300km). Examples of such separation seen by the NPL scientists are shown in Fig. 7.

### INTERNATIONAL GEOPHYSICAL YEAR (IGY): 1957-58

When the IGY came, a very sound base on radio science had been created. The laboratory of Professor S K Mitra at Calcutta was in full swing.

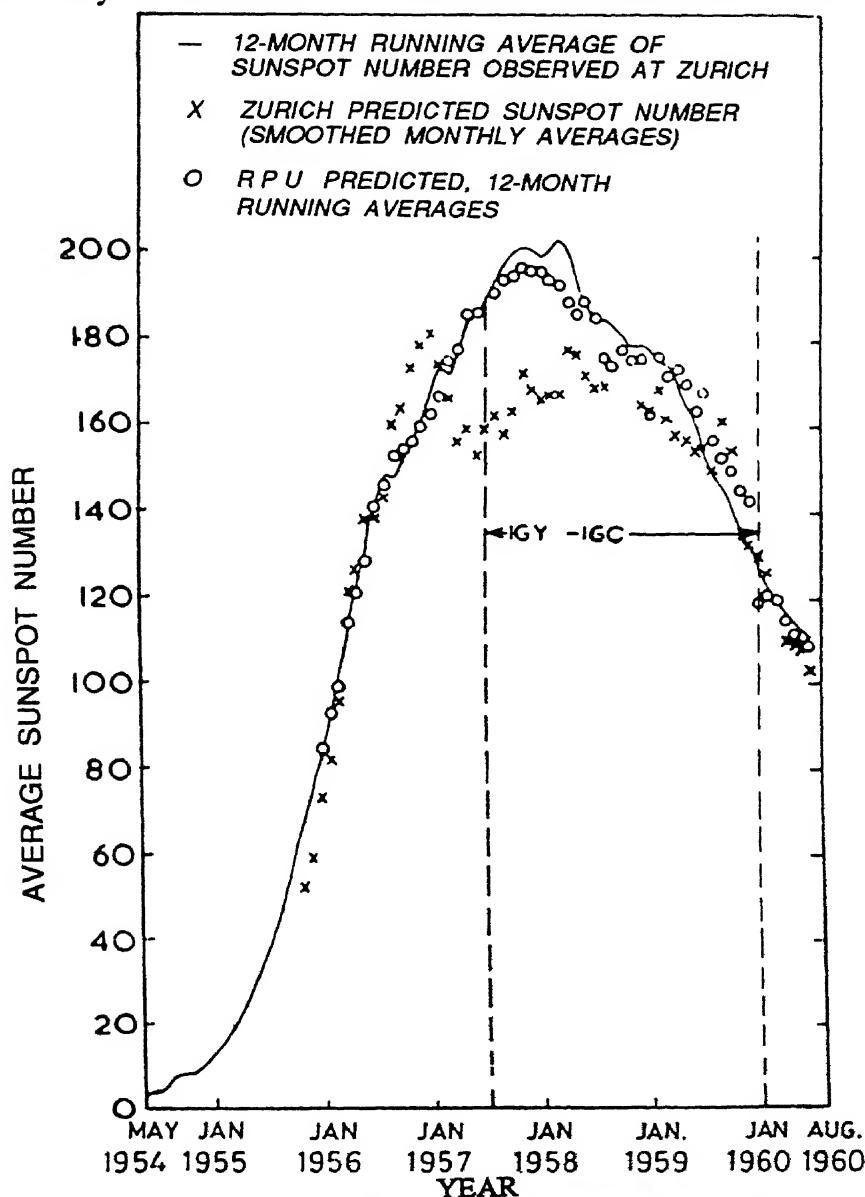


FIG. 8. The highly successful NPL predictions for sunspot activity during the IGY vis-a-vis predictions from Zurich

Physical Research Laboratory had been established at Ahmedabad under the leadership of Sarabhai and Ramanathan. All India Radio had organised a major programme of establishing and running a network of ionosphere recorders at Delhi, Madras, Tiruchirapalli, Kodaikanal and Bombay and the Radio Propagation Unit has just been formed at NPL under A P Mitra. The latter started the job of coordinating the ionospheric data in India and also the work of prediction of solar activity needed for radio communication predictions. This has continued now for some 40 years and has proved remarkably good. During the IGY, which happened to coincide with the largest sunspot maximum in recorded history, the prediction system introduced in Delhi fared better than the Zurich prediction which was then in wide use.

This is shown in Fig. 8. Two new techniques were put into operation: riometers and S N Mitra's ionospheric drift technique.

The drift technique was developed initially in the Cavendish Laboratory in Cambridge in 1946 (in which S N Mitra was a collaborator) and was introduced in India during the IGY in a major way (involving stations at Ahmedabad, Delhi and Waltair)<sup>31</sup>.

Radio techniques of a wide variety were commissioned for the detection of solar flares: these included the use of atmospherics, cosmic radio noise, vertical incidence ionosondes, radio emission from the Sun. We have had thus one of most extensive radio patrols of flares undertaken anywhere in the world. An example is shown in Fig. 9. This formed the basis of A P Mitra's book on *Ionospheric Effects of Solar Flares* to be brought out some two decades later and providing the first look into the chemical changes in the Ionosphere during flare events.

India's special position near the geomagnetic equator was utilised. The presence and the complex nature of the electrojet straddling the geomagnetic equator and the curious phenomenon of ionisation peaking not at geomagnetic equator but at the dip value of about  $30^\circ$  away from it were already known. Ionospheric stations were carefully selected and located to take advantage of this phenomenon. A network of 11 ionospheric stations (Fig. 10) were organised ranging over latitudes from  $28^\circ 38' \text{N}$  to  $08^\circ 29' \text{N}$ . Of these eleven stations as many as five were in the anomalous equatorial zone. Trivandrum was located almost exactly on the magnetic equator, Kodaikanal on the fringe of the electrojet and Tiruchirapally in the middle of it. In north, Ahmedabad and Calcutta were located at the peak of the Appleton Anomaly and they were (excepting

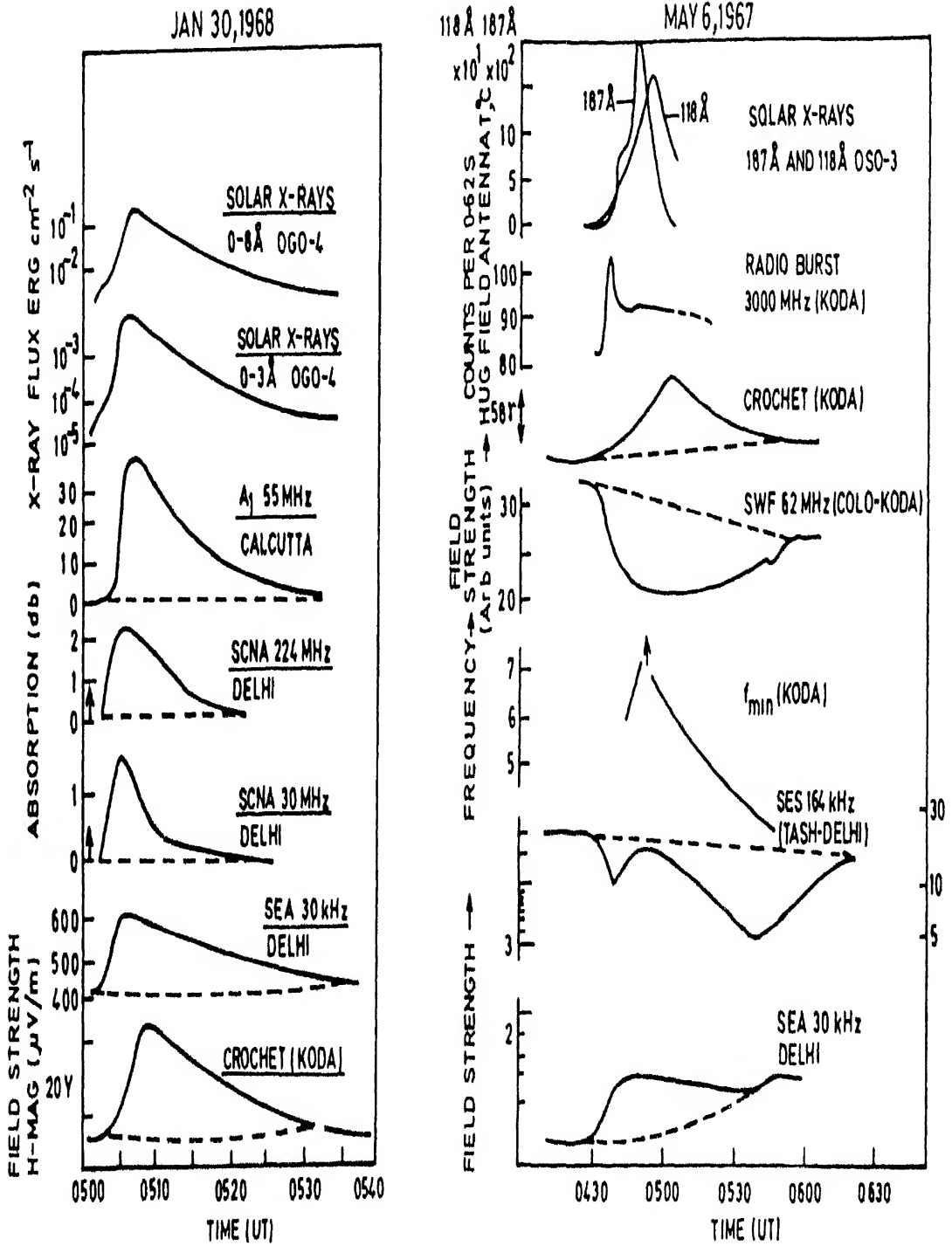


FIG. 9 Examples of two cases of solar flares recorded by a variety of techniques at the NPL, New Delhi (after A P Mitra)

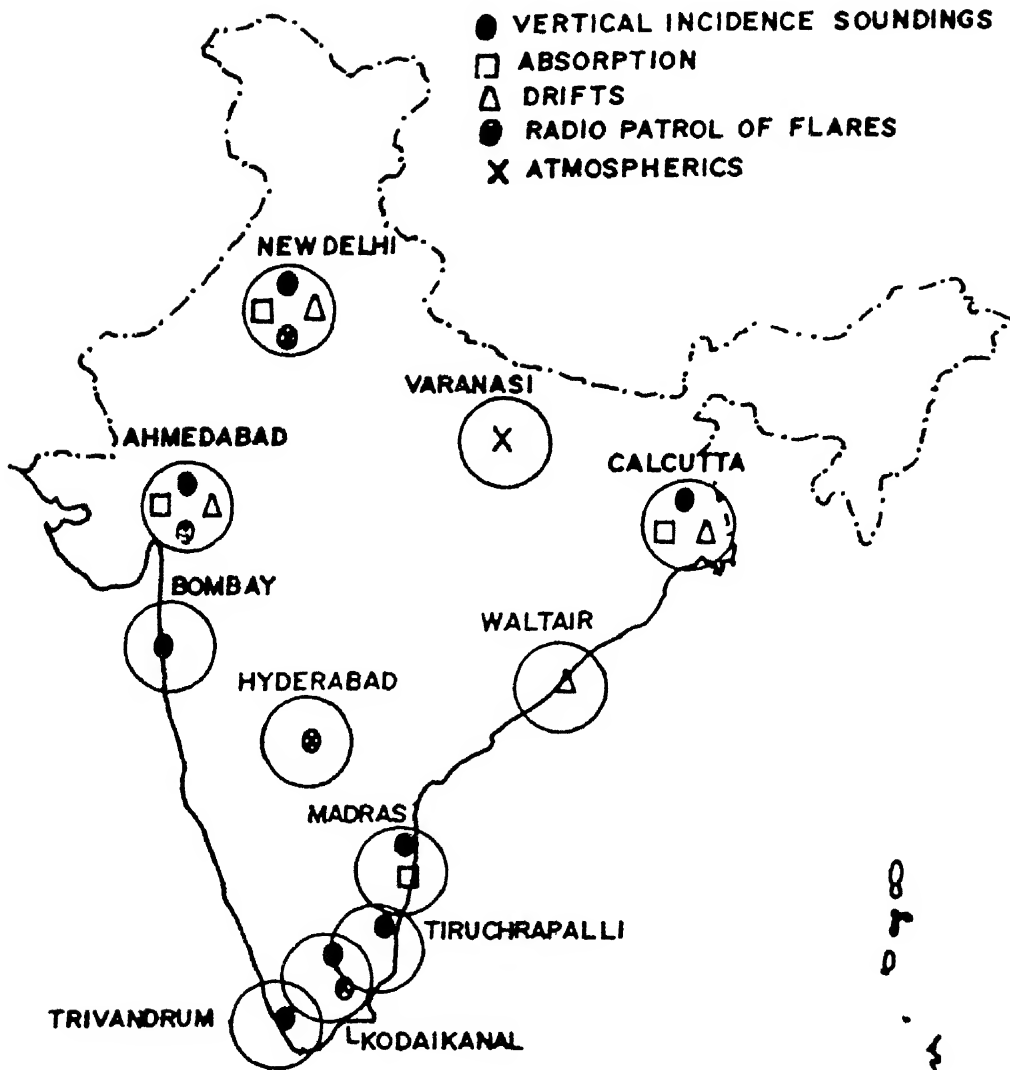


FIG. 10 IGY stations in India for ionosphere and for radio monitoring of the Sun

Calcutta) all nearly around the same longitude:  $75^{\circ}\text{E}$  — an unusual advantage that was immediately recognised.

The IGY provided, for the first time, an organized entry of Indian radio science into international arena. Much of the coherence in the nationally coordinated programmes that has now become possible in ionospheric and magnetospheric physics and in radiocommunication problems is a consequence of this beginning of organized research in India.



## SPACE RADIO RESEARCH

Space radio research began as early as 1957, when the first Satellite was launched. Radio scientists at NPL and elsewhere started recording the telemetry transmission from the satellite. Serious radio beacon observations, however, began with the satellite COSMOS 5 at frequencies of 20, 40 & 41 MHz<sup>9</sup>.

The most important event, however, was the beginning of rocket launching facility from Thumba in 1963. Although the first experiment was the release of sodium vapour for the measurement of upper atmospheric drifts, subsequent experiments included techniques of radio science interest. One of the most successful series of experiments of this kind involved the measurement of ionisation and plasma instability using Langmuir probes of special design by Satyaprakash and his colleagues at PRL<sup>10</sup>. It was possible to measure amplitudes and spectra of these irregularities in a wide range of scale sizes. These measurements were the first of its kind. A number of other rocket experiments quickly followed: use of resonance probe, propagation experiments with transmitters at the ground and receivers in the rocket (extensively used by Somayajulu of NPL). A novel experiment that Indian scientists conceived and successfully carried out was the placement of a riometer receiver in the rocket<sup>11</sup>. Use of this rocket facility has continued: extensive use was made particularly during the Indian Middle Atmosphere Programme (see later).

A regular programme began in 1962 when the NPL started receiving 20MHz radio beacon signals from the Russian orbiting satellite COSMOS 5. Faraday rotation of these signals was recorded as it passed through the ionosphere for about a year until the beacon transmission ceased. These measurements allowed modelling of the outer ionospheric ionization distribution and also indicated the nature of transformation of  $O^+$  to  $He^+$  ions<sup>9,12</sup>. The satellite radio beacon efforts were to remain a major and critical part of ionospheric activity for the next two decades. One of the most important efforts in the mid 70s was the programme connected with the ATS-6 satellite (Fig.11). There were two major programmes: (1) one relating to direct reception of TV signals from about 2330 villages in six widely separated direct reception clusters and (2) a planned network of satellite radio beacon receiving equipment operating in different parts of India at one or more designated frequencies: 40, 140, 360 and 860MHz. The direct broadcast satellite programme (known as SITE) was an unique experiment conducted jointly by ISRO and NASA for family planning,

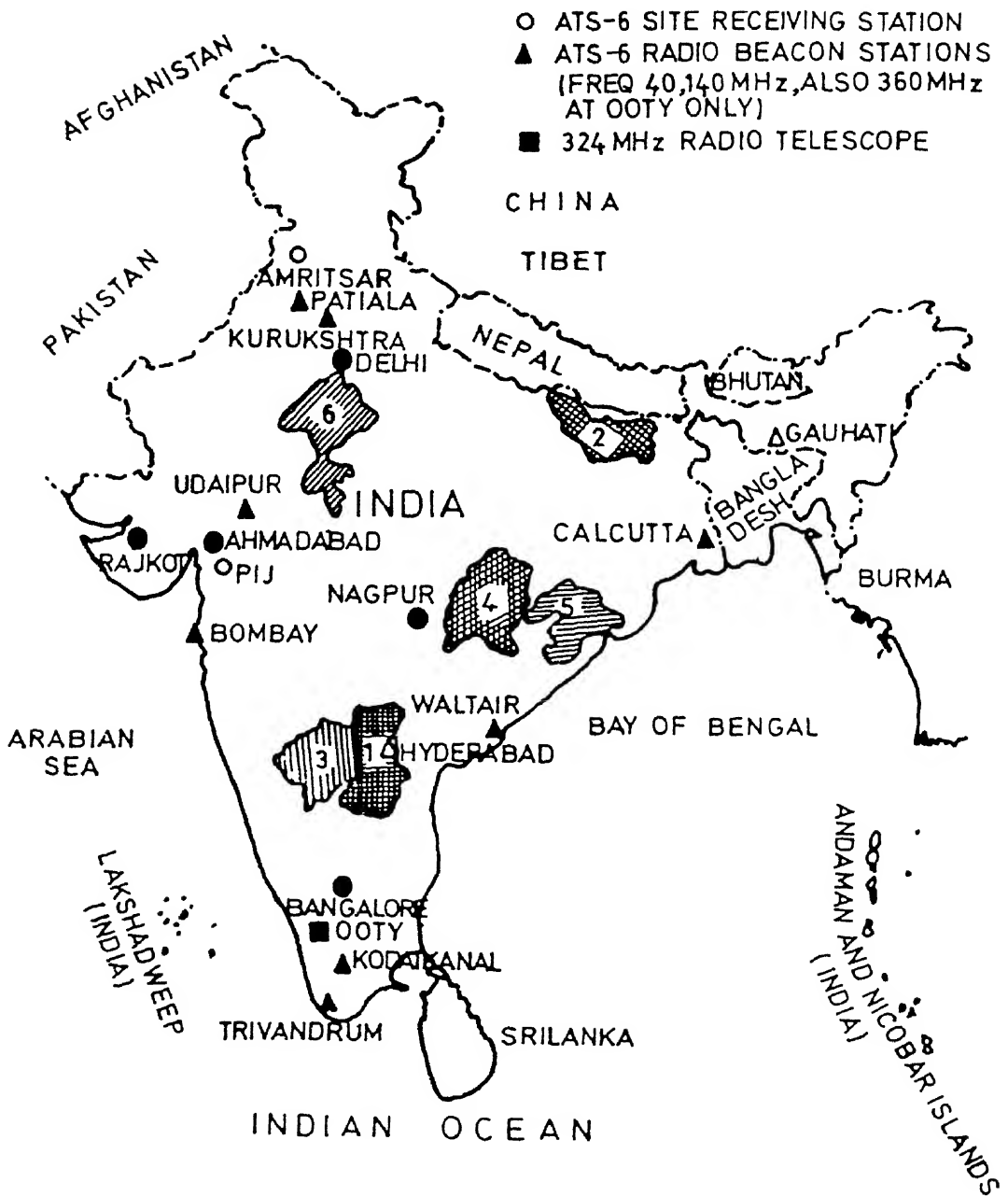


FIG 11 Site and satellite radio beacon experiments with ATS-6 satellite

agriculture, health and hygiene. The prime earth station was located at Ahmedabad, linked to a TV studio which in turn was connected to a VHF transmitter at *Pij*, 50km from Ahmedabad. Programmes were sent to the satellite on 6GHz band and were then retransmitted by the satellite at 860MHz with a 30ft antenna having a pointing accuracy of 0.1°. Other earth stations were at Delhi and Amritsar, connected to VHF transmitters

of AIR. Each village receiving system consisted of: (a) a chicken mesh dish of 10 ft diameter; and (b) a front end converter which converted the VHF signals to video and audio. Both were designed and developed by ISRO and manufactured by ECIL.

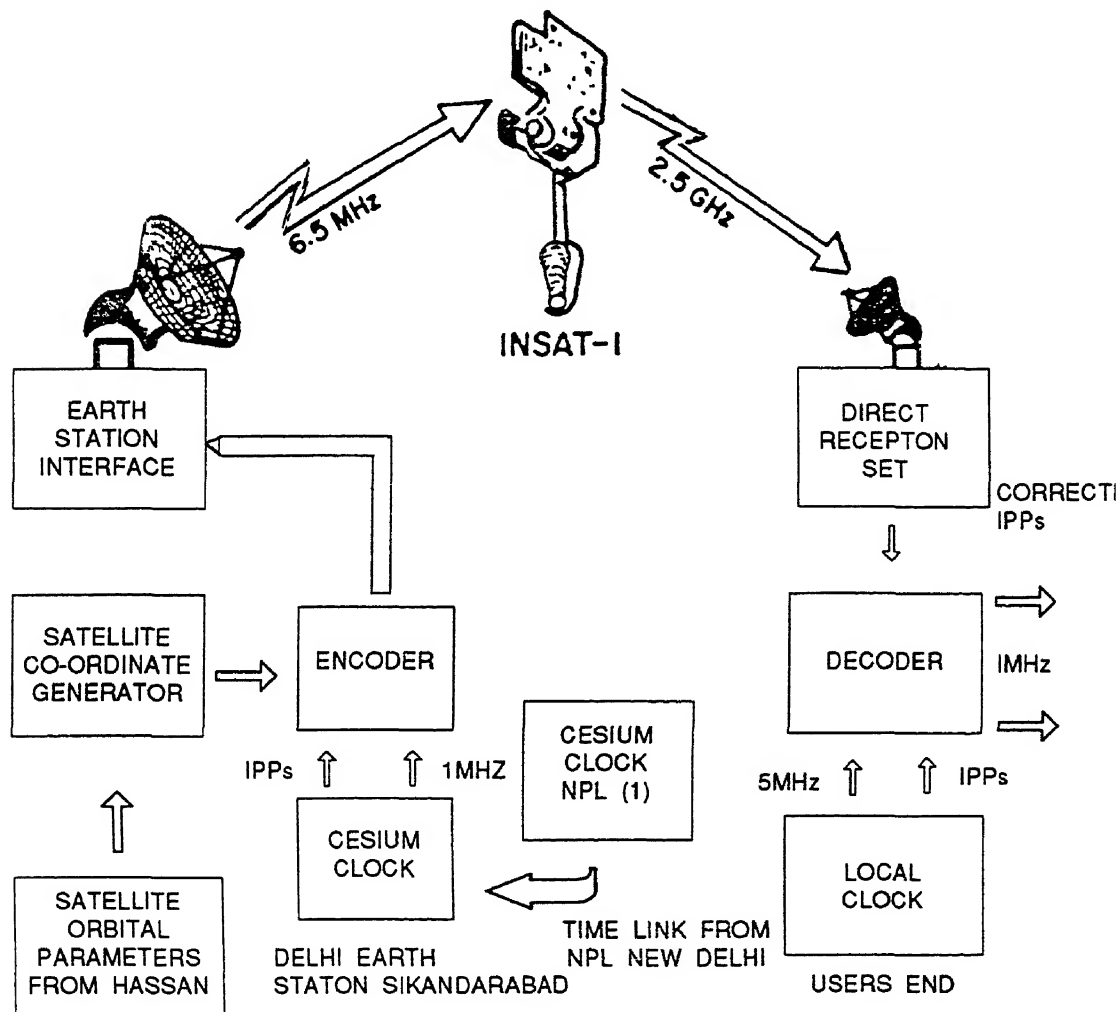


FIG. 12 Set up of satellite time transfer in India, as operated by the NPL (courtesy: B S Mathur and group)

The earth stations were also designed and assembled within the country. The experiment was successful both scientifically and socially. One of the many examples of its success was the Republic Day telecast on January 26, 1976. On this day the satellite network included all Doordarshan TV stations. The result was that for the first time people in Delhi, Calcutta, Bombay, Poona, Srinagar, Lucknow, Khera and 2330 villages saw the parade in real time<sup>27</sup>.

While the SITE programme was operationally an exciting adventure, the second aspect of the ATS-6 programme—the satellite beacon experiment (Fig. 11)—was also scientifically exciting. ATS-6 was the first *geostationary* satellite available (all previous studies were with low orbiting satellites) to a wide community of Indian ionospheric scientists; its coverage of frequencies extending from 40MHz to 860MHz offered scope for study of the frequency effects in scintillation in equatorial regions including those at the frequency of 860 MHz then considered too high for ionospheric effects. An extensive network of stations could be organised involving the National Physical Laboratory at Delhi, the Physical Research Laboratory at Ahmedabad, Vikram Sarabhai Space Centre at Thumba, the Radio Astronomy Centre at Ooty, and a number of Universities including Patiala, Kurukshetra, Delhi, Udaipur, Calcutta, Saurashtra and Andhra<sup>26</sup>. All these stations were located around 70°E longitude and the coverage extended from geomagnetic equator to about 27°N geomagnetic latitude—an usually attractive range of latitudes for the study of low latitude phenomena, especially ionospheric scintillations that are known to be particularly severe in the equatorial belt. Another special feature was the availability of the Ooty telescope for measurement of scintillation of selected radio stars in the same time frame. This was one of the most extensive chains.

For the study of equatorial ionospheric scintillations, the Indian observations continue to provide the most significant base data. Many results are available: predictions, spatial and temporal, have also been made, although here input information is not adequate. Space systems were also brought in late 70s to provide a major new dimension in time and frequency dissemination and calibration. This came in two ways: (a) through replacement of old quartz clocks by cesium atomic clocks and (b) from the use of satellites. The French-German satellite *Symphonie* was used to synchronize (a) two atomic clocks located in Madras and Ahmedabad and (b) externally between NPL (India) and PTB (West Germany), the latter with a precision of better than 10ns. This satellite was also used for one-way coded time transmission and time transfer *via* direct TV broadcasts. Use was made also of the Indian experimental satellite APPLE, INTELSAT-IV and Indian domestic satellite INSAT 1-B and time intercomparison was carried out between NPL and VNIIF-TRI (USSR) *via* direct TV broadcast from Russian satellite STATIONER-6 in December 1985.

Time transfers through INSAT series as a regular system have been initiated by the NPL recently (currently through INSAT 1-D) (Fig. 12) with the transmission setup located at the earth station in Sikandrabad, UP and transferring time with an accuracy  $\sim 10\mu\text{s}$ . There is increasing use of the service amongst specialised users. In addition, a Global Positioning Satellite Receiver (GPS) has been installed at NPL and is being operated as part of international time intercomparison work.

Remote sensing with the satellite BHASKARA I and II was the next major milestone in radio research in India. There were two nearly identical low orbiting satellites at heights around 520km and inclination 51, carrying microwave radiometers (SAMIR).

BHASKARA I operated from June 19, 1979 to March 1981; BHASKARA II was launched in November 1981. There were two radiometers in BHASKARA I (19.35 and 22.235GHz) and three in BHASKARA II (19.35, 22.335 and 31.40GHz). With these three frequencies different types of information could be obtained through the 22GHz radiometer atmospheric water vapour content, and through all frequencies information on "sea state" and rainfall rates. The 22GHz and 31 GHz radiometers thus provided dual functions: (a) gave atmospheric corrections for sea surface studies; and (b) at the same time provided information on atmospheric water vapour and liquid water content. Observations have been carried out principally by Space Applications Centre, Ahmedabad, National Remote Sensing Agency, Hyderabad; and more recently NIO has been involved. Other developments currently in progress include. 8-18GHz FM — CW Scatterometers, Side-looking Airborne Radar (SLAR), Air-borne Synthetic Aperture Radar (SAR); multifrequency scanning radiometer. The X-band SLAR developed by National Remote Sensing Agency has been flight tested in an aircraft. Both SLAR and SAR are expected to be in observational use soon. There are also plans to use microwave remote sensing from the European Satellite ERS-1. All these efforts are directed towards having radio remote sensing as a supplementary input to optical remote sensing systems.

### THE FERRITE STORY

In 1963, the Government of India appointed under the Chairmanship of H J Bhabha an Electronics Committee to prepare a 10-year development plan for electronics (BARC Committee and Electronics 1966).

The Committee noted that since the development cycles in several areas were of the order of three years, intense and continuous research was vital, adequate test facility (then woefully inadequate) essential, and production of components (backbone of electronic industry) of adequate quantity and quality, a matter of highest priority.

The deliberations of the Bhabha Committee gave an impetus to the technology development then underway in NPL (under Ramamurti and Ganapati) on electronic components (especially mica and ceramic capacitors and ferrites) and the efforts on instrumentation development in BARC. The ferrite story is one of the early successes in indigenous development of electronic components. It held its own against competitions of world class technologies abroad (such as those of Phillips). Its laboratory-to-plant progress is shown below:

#### LAB TO PLANT

LAB: NPL

PLANT: CEL

PRODUCT: PROFESSIONAL FERRITES

1972: 10 KG. BATCH PRODUCTION AT NPL

1974: TRANSFER OF KNOWHOW TO CEL

1977: SCALE-UP TO 100KG. BATCH

1984: MICROWAVE FERRITES FOR DEFENCE

1985: EXPANSION TO 200KG. BATCH

1987: EXPANSION PLAN FOR 500 TPA

#### **Adges: Defence/Academic Institutions Collaboration on Tropospheric Research**

Early 70s saw the beginning of intense troposcatter research in the country. There was already a sizable activity in radio meteorology. Indigenously designed radiosondes for upper air meteorological observations (which had started in the mid 40s) had been in operation for more than two decades. In the late 60s, they had been replaced by new audio-frequency modulated systems in the national network. Radars were also being extensively used: X-band radars had been introduced in the mid-50's for detection of thunderstorms and S-band radars in the mid 60s on the East and West coast for detection of cyclonic storms. Radio refractivity computations had also begun in early 60s, on the basis of radiosonde data. Anomalies had been seen in the VHF links such as those operated by the Aviation Department involving aircraft-to-ground and ground-to-aircraft communication of 108MHz. Anomalous TV signals had also been observed and examined. But the main impetus came from:

(a) the desire in India to establish transhorizon troposcatter links; and (b) the increasing sophistication and requirement of radar operation. Consequently, radar and troposcatter system research was initiated at the 5 IITs, the Indian Institute of Science in Bangalore, the Roorkee University and the National Physical Laboratory. A wide variety of activities were initiated: phased array radars, establishment of troposcatter links, monitoring of tropospheric radio parameters through radiosondes, introduction of new techniques of monitoring etc.

One of the most important works at this time was the preparation of an Atlas of tropospheric radio refractivity over the Indian sub-continent by the National Physical Laboratory and the India Meteorological Department<sup>13</sup>. This gave surface measurements of pressure, temperature and humidity recorded at 36 meteorological stations for a period of 5 years (1959-63) and radiosonde observations over 16 stations for a period of 4 years (1968-71). The Atlas was subsequently revised to include observations of 32 radiosonde stations (instead of 16 in earlier Atlas). Since the user requirements, because of increased system design, demanded new types of data such as those on turbulence, structure constant, better coverage over oceans, subrefraction and superrefraction occurrence frequencies, the new Atlas included estimates of  $C_N^2$ , subrefraction percentages to aid fresnel zone clearance, estimates of refractivity profiles over oceans using Monex data<sup>14</sup>. The Atlas which is continuously updated has been recognized as a basic reference document for all troposcatter and radar system operation in India. A representative distribution of occurrence of radio ducts is shown in Fig. 13.

Several experimental links were set up. One was the troposcatter link at 2.1GHz set up between Nainital and Kanpur using a 1KW transmitter located at Nainital and receiver at Kanpur. The distance between the two was 330km. For characterising the radio refractivity conditions over the path a microwave refractometer loaned by the USA was used in an aircraft in the initial stages. Subsequently two solid state microwave refractometers were built by joint efforts of NPL and DEAL (DRDO) and flown in a CESSNA aircraft in several sorties to collect data. The refractometers operate on 9.34GHz and have a sensitivity of 0.1N and 6 $\mu$ s response time. Several links were set up around 2GHz in Northern India for specific operational purposes. One particular path under intensive study was the 158km Delhi-Pilani path.<sup>15</sup> These were also used for scientific studies. In addition, in late seventies and early eighties,

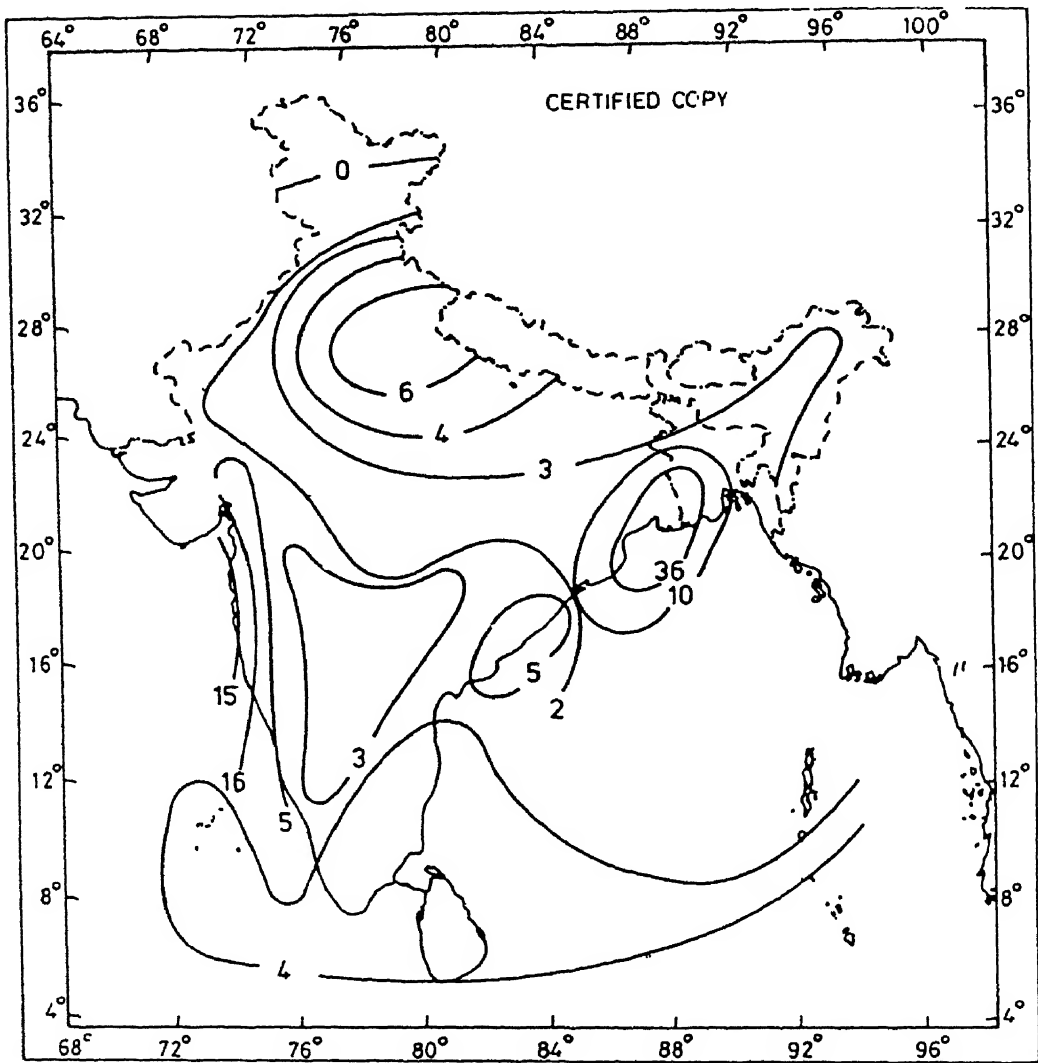


FIG. 13 Contours of radio duct occurrence probability over the Indian subcontinent (May, 1200 GMT) (from the Atlas of Radio Refractivity)

propagation parameters were extensively studied using operational P&T links and in the VHF with TV transmitter.

New techniques of tropospheric monitoring were introduced: one such technique was the sodar. Fig. 14 gives a photograph of one of the earliest sodar antennae located on the grounds of the NPL. The sodar is essentially an acoustic monitoring of atmospheric inhomogeneities; it locates and measures the intensity of thermal and velocity inhomogeneities—an information which can be used to infer atmospheric processes and their time histories. The sodar<sup>16</sup> that was conceived, designed, fabricated and set up in NPL operated at a frequency of 2kHz,



had a height range of about 600 meters and initially used a microwave dish (luckily available as spare in the NPL), acoustic sources embedded in the ground with sandbags as shields and a facsimile recorder graciously loaned by IMD. This improvised model (costing only a few thousand rupees but sufficient to prove its potential) had been continuously improved over the years with different antenna systems: horns, parabolic dishes and shields of various types such as polyurethane foam. The sodar provided the first detailed mapping of the ducting conditions. To calibrate the sodar it was operated simultaneously with radiosondes in Ayanagar near Delhi. Subsequently a sodar was located by NPL at BARC and a doppler sodar was set up at VSSC in Trivandrum. A radio acoustic system RASS is now being designed for installation at the MST site.

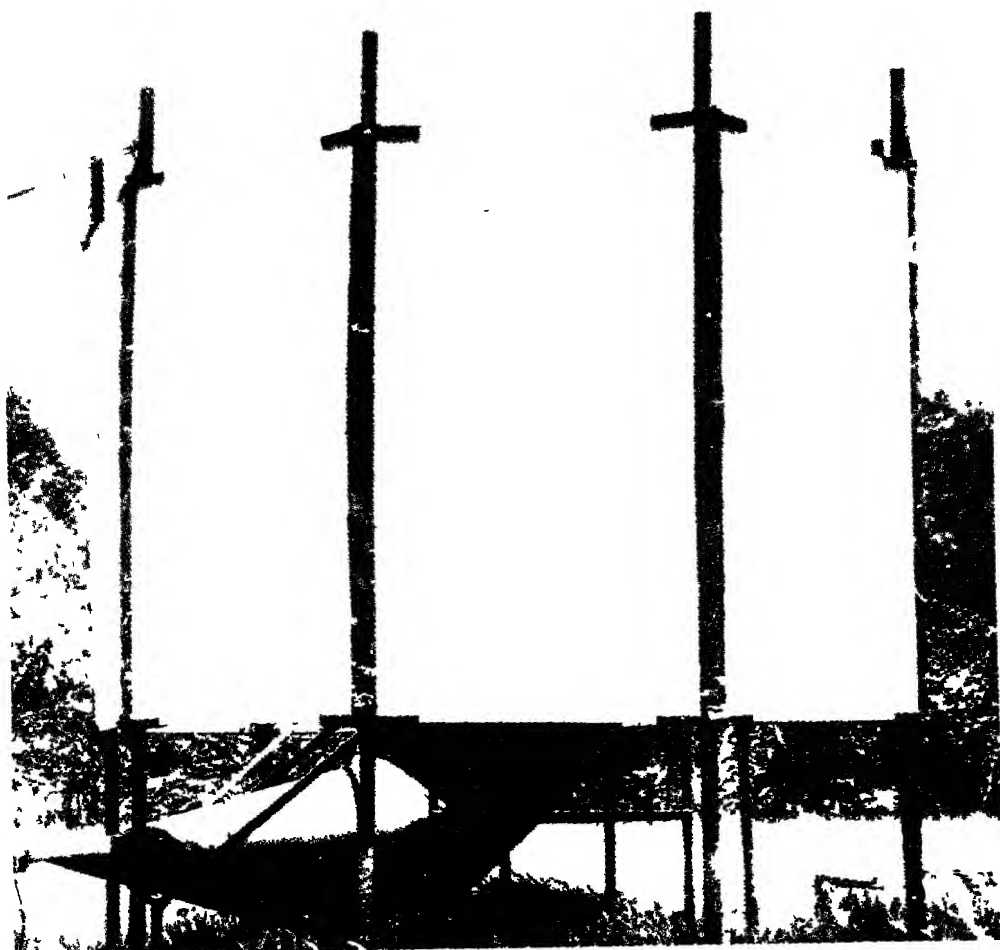


FIG. 14 One of the earliest SODAR set up fabricated at NPL (*courtesy: S P Singal*)

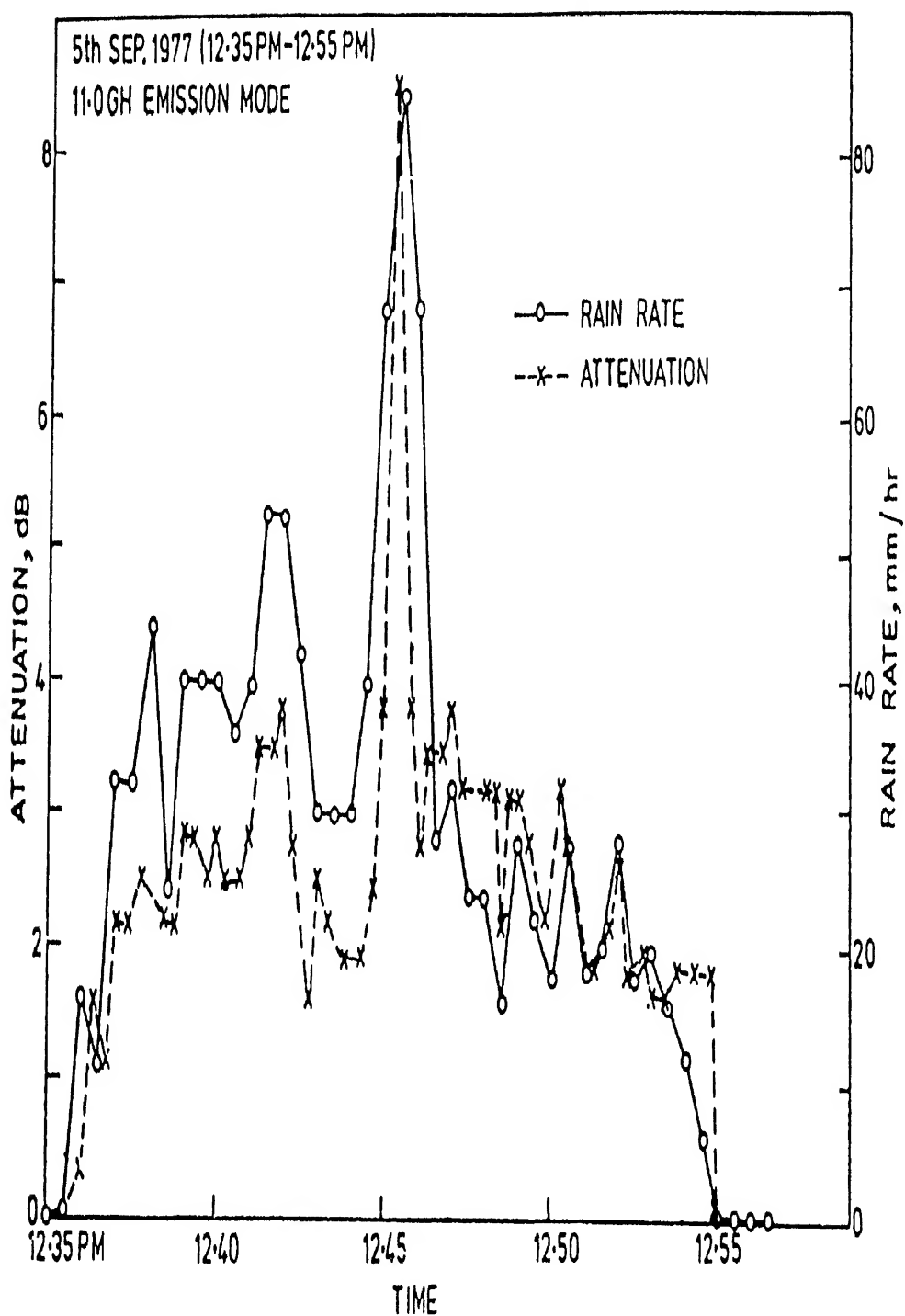


FIG. 15 The complex and highly variable structures in rain rate during a tropical rainstorm and the simultaneous attenuation surges in microwave region (after M K Raina)



FIG 16 The Ooty radiotelescope built by TIR (courtesy G Swarup)

The third technique was the use at the NPL of groundbased microwave radiometers designed and installed at 11, 18 and 22.3GHz to monitor rain attenuation and water vapour in the atmosphere<sup>17</sup>. The radiometers operated on both passive and active (the sun as source) modes. The rain attenuation measurements were made simultaneously with rain rate measurements with an electronic rain gauge built specifically for this purpose (by a young NPL scientist at NPL Teddington) and with time resolution of 10 seconds. Tropical rainstorms have many short duration surges; these produce simultaneous attenuation surges that must be taken into account in equipment design (Fig. 15). By combining troposcatter specific link observations at two frequencies with initial refractivity gradients measured with the radiosondes, a new tropospheric transmission loss prediction system was worked out by Mazumdar<sup>18</sup>. This has since been extensively used in the country. It forms the foundation of present day tropospheric designs in India.

### RADIO ASTRONOMY

The most spectacular facility of the seventies was the radio telescope (Fig. 16) built by TIFR at Ootacamund. Construction of the telescope was started in 1966 and was completed in 1970. It operates at one frequency band centred on 326.5MHz and has an effective collecting area of 8000m<sup>2</sup> (about three times that of the Jodrell Bank's 250ft dish) with a 530m long and 30m wide parabolic cylindrical antenna array. It was designed primarily for lunar occultation observations of weak (and thus very distant) extragalactic sources. An important feature of this telescope is fact that the long N-S axis is parallel to the earth's axis of rotation; this has been arranged by placing the telescope on a north-south slope (inclination to the horizontal of about 11°) equal to the latitude of the place. It thus allows tracking of a radio source for about 9.5h in hour angle by simply rotating the antenna system along its large axis. Since 1970, lunar occultations of over 1200 sources have been observed; with a median flux density of 0.6 Jy at 327MHz. The occultation survey provided much information on the structures of weak extragalactic radio sources. An important result was the establishment of angular size-flux density relationship. The major inference that measured angular sizes were systematically smaller than what one would expect from nearby strong sources provided a major (and independent) support of the big-bang model of the origin of the universe<sup>34</sup> (Fig. 17).

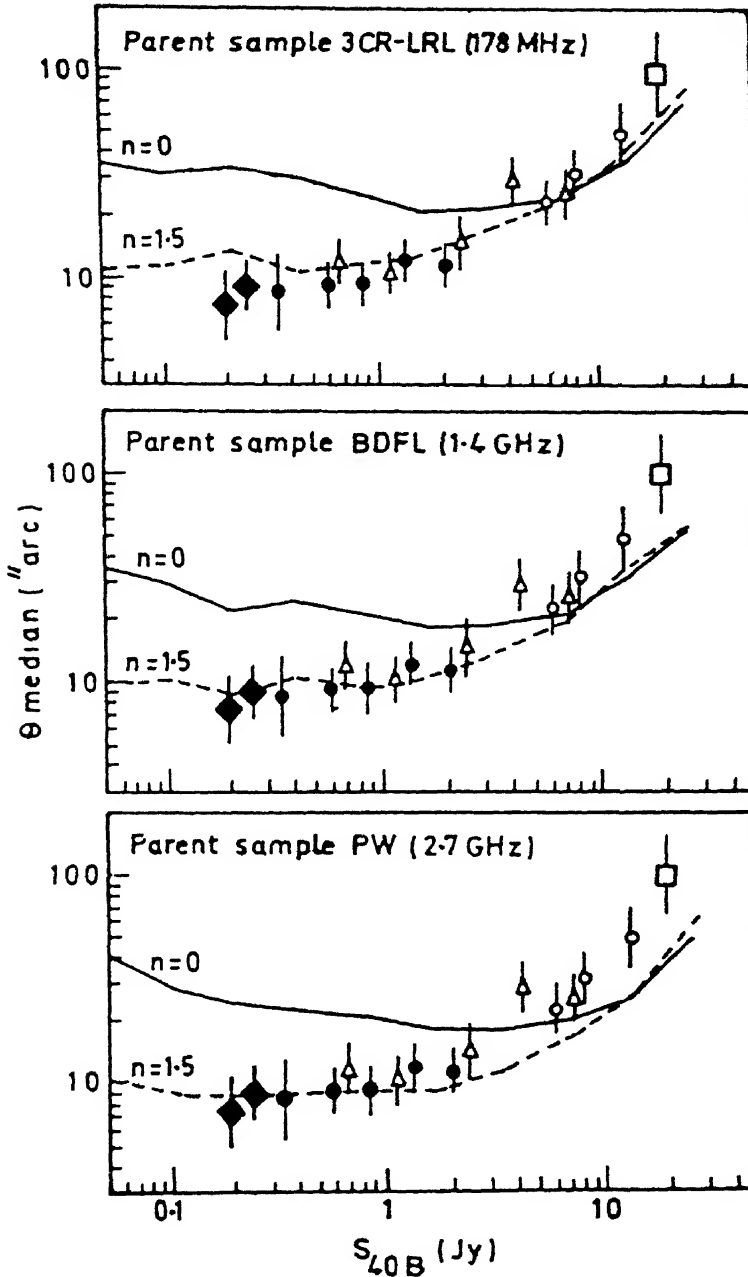


FIG. 17 Angular size-flux density relation for extragalactic sources as obtained with the Ooty radio telescope and predictions of an evolutionary model of radio luminosity function

In radio astronomy there were several new major developments in Ooty: the Ooty Synthesis Radio Telescope—the big cylinder was combined with several baby cylinders (22m X 9m), all steerable, to produce a 4km synthesis radio telescope (OSRT) having a resolution of  $40 \times 50$  arcsec at 327MHz and a signal to noise ratio of 5:1 for a 20mJy source for 8 hour

observing run. This is being used for detailed mapping of strong radio sources, nearby galaxies, clusters of galaxies, supernova remnants, H II regions and radio stars (an excellent summary has been given by Swarup)<sup>19</sup>.

A new activity initiated during this period concerned study of interplanetary scintillation (IPS)<sup>22</sup>; these occur when radiowaves from a distant compact radio source (such as a quasar) propagate through the interplanetary medium where these are scattered by irregularities in plasma density of scale sizes  $\sim 100$  km. Some early work on IPS was done with the Ooty telescope with interesting results: revelation of a deficit of compact scintillating sources around the inner regions of our galaxy, estimation of scale sizes of irregularities in the interplanetary medium and possible identification of plasma irregularities in the tail of comet Kohoutek<sup>20</sup>. An equipment facility specifically devoted to IPS observations has been installed by the PRL. This consists of a three station array on a frequency of 103MHz at Thaltej (near Ahmedabad), Surat and Rajkot. The Thaltej station has a collecting area of 10,000m<sup>2</sup> and Rajkot and Surat, each around 5000m<sup>2</sup>. All the three telescopes have crystal-controlled clocks assuring synchronization in operation with relative time accuracy of a few milliseconds. Several radio sources have already been used. Daily solar wind is being monitored essentially regularly. At Thaltej, day to day changes by a few arc minutes were observed at 103MHz with certain radio sources. There was uncertainty whether the necessary largescale gradients occur in the Earth's ionosphere or the interplanetary medium. This matter was resolved through simultaneous observation of a radio source with IPS telescope and observations of signals from the geostationary satellite Fleetsat with a satellite beacon receiver at 244MHz in an interferometer mode with antenna separated by 100m. Another interesting result was the observations on scintillations observed on quasar PKS 2314+03 during 18-20 December 1985 as it was being occulted by the plasma tail of Halley's Comet. There was enhancement in scintillations indicating existence of turbulence in the cometary tail. A rough estimate of r.m.s. electron density fluctuation in the plasma tail was 4.4cm<sup>-3</sup> at 0.12 AU and 1.3cm<sup>-3</sup> at 0.18 AU. Similar observations have also been made more recently (1-2 May 1987) with radio sources PKS 0606-796 and PKS 0637-752 occulted by Comet Wilson and 1 May and 2 May 1987.

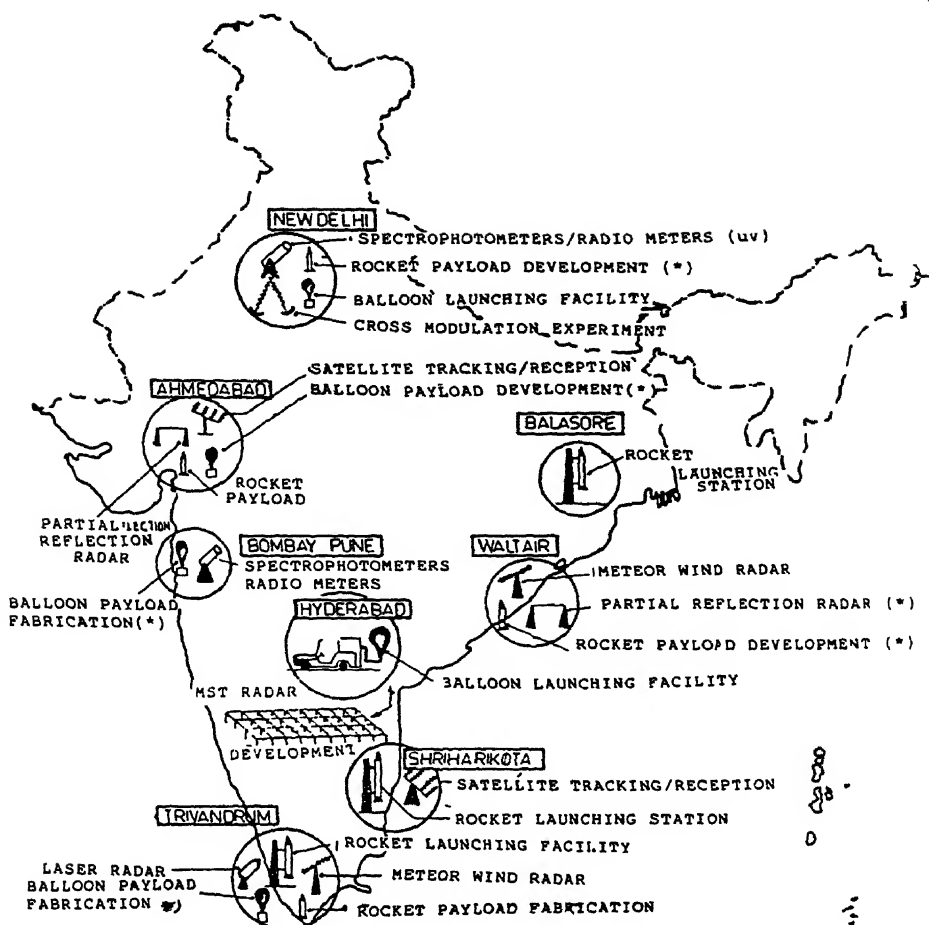


FIG 18 Facilities and Institutions participating in the Indian Middle Atmosphere Programme

Another major new facility which has recently been made operational is a mm wave radio telescope with 10.4 metre antenna operated in the range 22-110 GHz; a facility of the Raman Research Institute. Two types of programmes are planned: the first to determine the continuum flux from quasars and compact nuclei of distant galaxies; the second objective is molecular lines spectroscopy. An interesting application has been the use of room temperature receiver operating around 110GHz for obtaining profiles of atmospheric ozone: preliminary results offer promise.

An interesting experiment performed in December 1983 for the first time in India involved very long baseline interferometry (VLBI) in which the participating observatories were Ooty (India), Jodrell Bank (UK), Westerbork (Holland), Effelsberg (FRG), Torun (Poland) and Crimea (USSR)<sup>21</sup>. The observations continued for 10 days and tens of sources (galaxies, quasars etc.) were monitored. A crucial component of the

experiment was high level clock synchronization. This was achieved with two rubidium clocks provided by the NPL and the P&T. The clocks were linked to the primary NPL cesium clock, with a synchronization of  $10\mu\text{s}$ . The corresponding frequency synchronization uncertainty was of the order of  $10\text{ns}$  over a few hours. A number of flat-spectrum radio sources were observed: the baselines ranged from  $0.2 \times 10^6$  to  $7 \times 10^6$  wavelength.

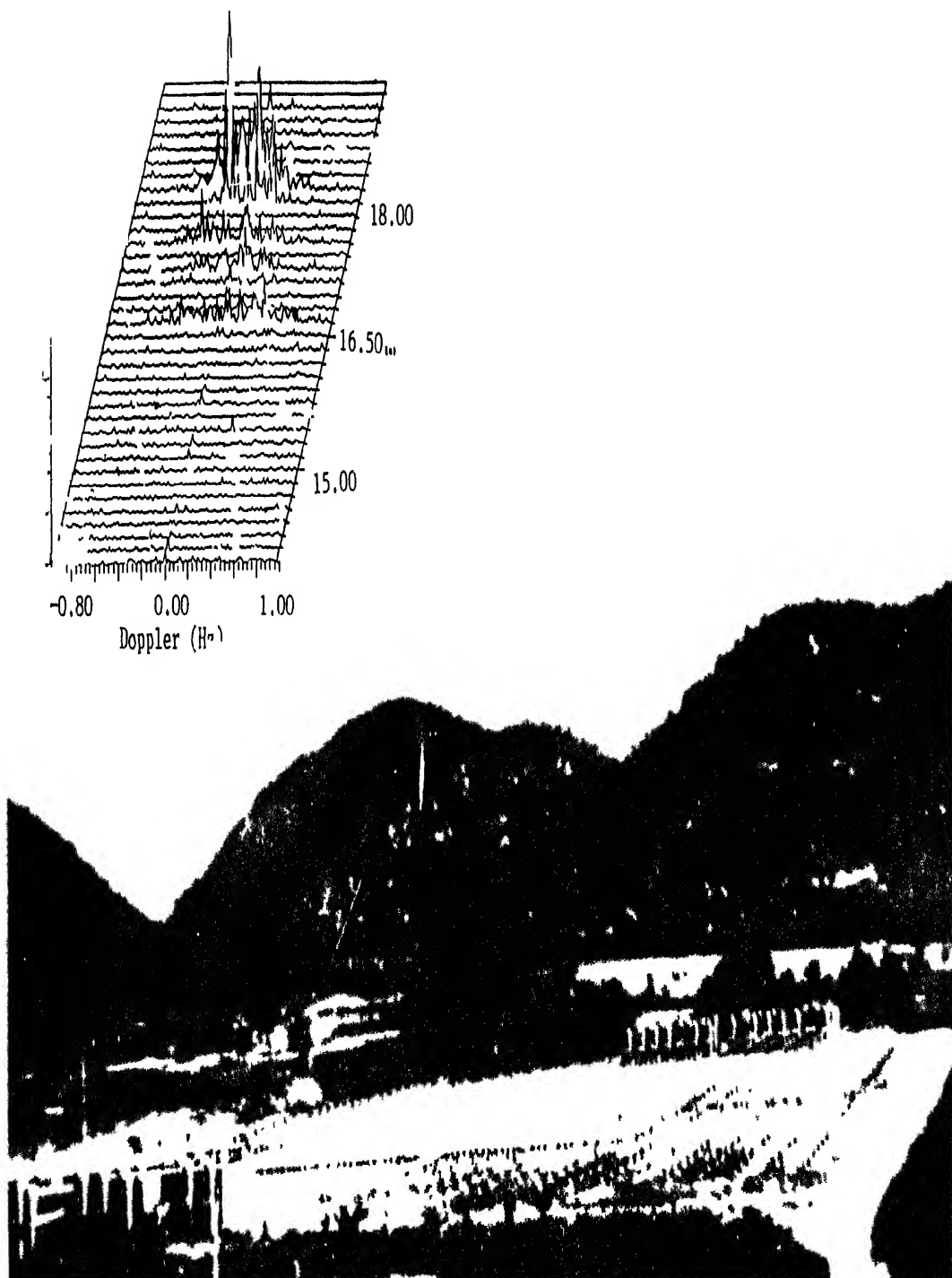
### THE MIDDLE ATMOSPHERE PROGRAMME

In 1982, the Indian Middle Atmosphere Programme (IMAP) was taken up, as a part of the International Programme, to study the region between 15 to 85km with a multiplicity of techniques: optical, radio, acoustic and with sensors located at the ground, in balloons and in rockets (Fig. 18). The programme originally scheduled till December 31, 1985, was later extended to March 1989. There have been three aims: (a) to examine the impact from man's activities, (b) the role played by the middle atmosphere in determining climate and climate changes, (c) to examine processes by which the sun, acting through the middle atmosphere, may be able to affect weather. For this programme a large number of major radio techniques were commissioned: meteor radars operating around 50MHz at Waltair and Thumba monitoring winds around 80-100km; a partial reflection Radar at metre wavelengths at Ahmedabad for mesospheric ionization and winds; a pulsed Doppler VHF backscatter radar at Thumba operating on 54.95MHz to obtain signal strength and doppler frequency spectrum of electrojet echoes; rocket and balloon borne systems to monitor middle atmospheric ionization and electric field (rockets were launched from Thumba, SHAR and Balasore; balloons from Hyderabad) and groundbased microwave radiometers to monitor atmospheric water vapour and its role in microwave propagation problems. The distribution of IMAP facilities is given in Fig. 18. The programme is the most extensive since IGY in the area of aeronomy, and its major strength is the introduction of campaigns with specific objectives (e.g. campaigns on stratospheric ionization, aerosols, dynamics, radiation) in which a variety of techniques were brought into operation in a carefully planned mode. As a result reference models for the Indian Middle Atmosphere are emerging.

### THE FUTURE

In the immediate future three major new facilities are planned: (1) MST Radar (Mesosphere-Stratosphere-Troposphere Radar), (2) GMRT (Giant





19 MST radar near Tirupati

Inset . Tropopause observations with 150m resolution (courtesy MST Radar Facility)

Metrewave Radio Telescope) and (3) SROSS satellite carrying on Ionospheric payload.

### *MST Radar*

The MST Radar is a high power, pulsed doppler radar operating in the VHF, capable of exploring the atmosphere from surface to about 90km, covering the troposphere, the stratosphere and the mesosphere. It is a joint effort of scientists from a number of national laboratories and Universities. It is a new generation radar of immense potential for continuous atmospheric probing in a three-dimensional form. Most MST radars are incoherent scatter radars modified for the purpose of monitoring heights below 100km: amongst the several exceptions the most noteworthy is the one recently established in Japan (the MU radar). The Indian MST radar is being installed at a place called Mittagadanki (13°N 79°E) near Tirupati not far from the rocket range SHAR—an added advantage providing possibilities of intercomparison and intercalibration. The radar was commissioned on ST mode on 29 October 1990 and operates on a frequency of 53MHz with peak power of 2.5MW and an average power of 60kW. There are 32 transmitter modules, with output powers in the range of 100KW to 20KW with a duty ratio of 2.5%. The antenna system is an array of 1024 yagi elements spread over an area of 136mX136m with an effective area of  $1.288 \times 10^4 \text{m}^2$  and a peak power aperture product of  $3 \times 10^{10} \text{W.m}^2$ . Five positions of beam have been selected: Zenith, 20 E—W and 20 N—S. A major point to note is that the scientific design has been completed entirely by Indian scientists and the fabrication has been undertaken by an Indian Group, SAMEER. When this radar is operational in its full MST form (expected by December 1992) it would be possible to monitor continuously atmospheric winds from close to ground to around 90km with a gap of around 35-55km.

It is a national facility providing major opportunities for atmospheric research to Indian aeronomers. A photograph of the antenna array and associated facilities is given in Fig. 19.

The Indian radar is one of the few similar radars currently in existence or planned. Global distribution of ST/MST radars is shown below:

*Global Distribution of ST/MST Radars*

| Facility Location    | Location    | Operating Frequency<br>(MHz) | Peak Power<br>Aperture<br>Produce $\text{Wm}^2$ |
|----------------------|-------------|------------------------------|---|
| Jicamarca, Peru      | 12°S, 72°W  | 49.9                         | $3.2 \times 10^{11}$                            |
| Pocker Flat, USA     | 65°N, 147°W | 49.9                         | $2.56 \times 10^{11}$                           |
| Mu, Japan            | 35°N, 136°E | 46.5                         | $8.33 \times 10^9$                              |
| Arecibo, Puerto Rico | 19°N, 67°W  | 46.8                         | $2.4 \times 10^9$                               |
| Sousy, Germany       | 52°N, 10°E  | 53.5                         | $1.92 \times 10^9$                              |
| Tirupati, India      | 13°N, 79°E  | 53.0                         | $3.12 \times 10^{10}$                           |
| Aberystwyth, UK      | 52°N, 40°W  | 47.0                         | $1.25 \times 10^9$                              |
| Chung-li, Taiwan     | 25°N, 121°E | 52.0                         | $5 \times 10^8$                                 |

To many, MST radar heralds a new age to atmospheric physicists much the same way radiosondes were decades ago. In the USA the technique is beginning to be used operationally for tropospheric weather prediction. One major reason is that it is the only instrument that can observe under all-weather conditions. Nevertheless, its acceptance by meteorologists has been slow and reluctant; this has been the case in most countries including India. Part of the reason is that the object weather prediction-forecasting of large scale atmospheric flow patterns (characteristic length scales of 2000km or more), associated precipitation effects, periods longer than a day—do not, it seems, benefit much from what is the principal advantage of the radar: its capability of sampling the atmosphere at short time intervals. But then there are events occurring on much smaller spatial and time scales: satellite cloud imageries have shown wide range in characteristic dimensions from a few hundred kilometers to almost synoptic scale (>2000km), and typical life cycles of about 12 hours. The typical convective system could have the following characteristics:

|                       |       |
|-----------------------|-------|
| Horizontal resolution | 100km |
| Vertical resolution   | 1km   |
| Temporal resolution   | 1hr   |

In this, the MST radar provides an important observing platform. There are other exciting possibilities: capability to locate atmospheric fronts, ability to detect echoes from precipitation particles as much as from refractive index irregularities (shown by the Japanese radar at Shigaraki), real time use of wind profiles in now-casting in regular use for the Program for Regional Observing and Forecasting Services in the USA), ability to monitor gravity waves and turbulence induced by the momentum

flux, observations of thunderstorm reflectivities, the capability to monitor the behaviour of the tropopause breaks and to locate the breaks relative to fronts. In the Indian context, what advantages will accrue in the study and understanding of monsoon dynamics is not immediately clear, but one should be able to monitor changes in circulation patterns of interest. For the mesosphere there is the distinct advantage of monitoring ionization turbulence fields of spatial and time scales otherwise unavailable. On occasions of increased reflectivity as during solar flares it should be possible to follow changes in increased scattered fluxes as the flare progress: this can open up a totally new look at flare time ion chemistry. Such a study has already been made by the Illinois group with Urbana radar.

### GMRT

The second major facility under construction is the GMRT—a giant radio telescope to operate in *meter wavelengths*, being located at Khodad, around 80km north of Pune<sup>33-35</sup>. The frequency range is 38MHz to 1420MHz. GMRT will be world's largest aperture-synthesis array with three times the collecting area of USA's Very Large Array (VLA) and with eight times its sensitivity. The design of the system has gone through several considerations: it is now a combination of 30 steerable parabolic dishes, each of 45m diameter with a novel design (thin wire mesh: 20mm X 20mm, 15mm X 15mm, 10mm X 10mm; stretched mesh attached to Rope Strusses—SMART) (Fig. 20a) and a configuration outlined in Fig. 20b. 12 of the dishes are placed randomly in a compact central array of about 1km in size. The remaining 18 form a Y-shaped array with each arm consisting of 6 antennas. The maximum baseline length is 25km. These are frequencies that are influenced by the terrestrial ionosphere and consequently available to ionospheric physicists as a new tool. At the same time for radio astronomers the undesired ionospheric effects can be overcome through the use of recently developed methods of self-calibration. The total effective area is 30,000m<sup>2</sup>. A 5 arcsec resolution is aimed for, and the telescope is expected to be in operation soon.

What are the objectives of such a major venture and why meter waves? The guiding principle is that the astrophysical problems to be examined should be such that these are either exclusive to the meterwave region or provide complementary information to that available at shorter wavelengths. Firstly, it will offer a powerful instrument to detect and study millisecond pulsars to detect primordial background of gravitational

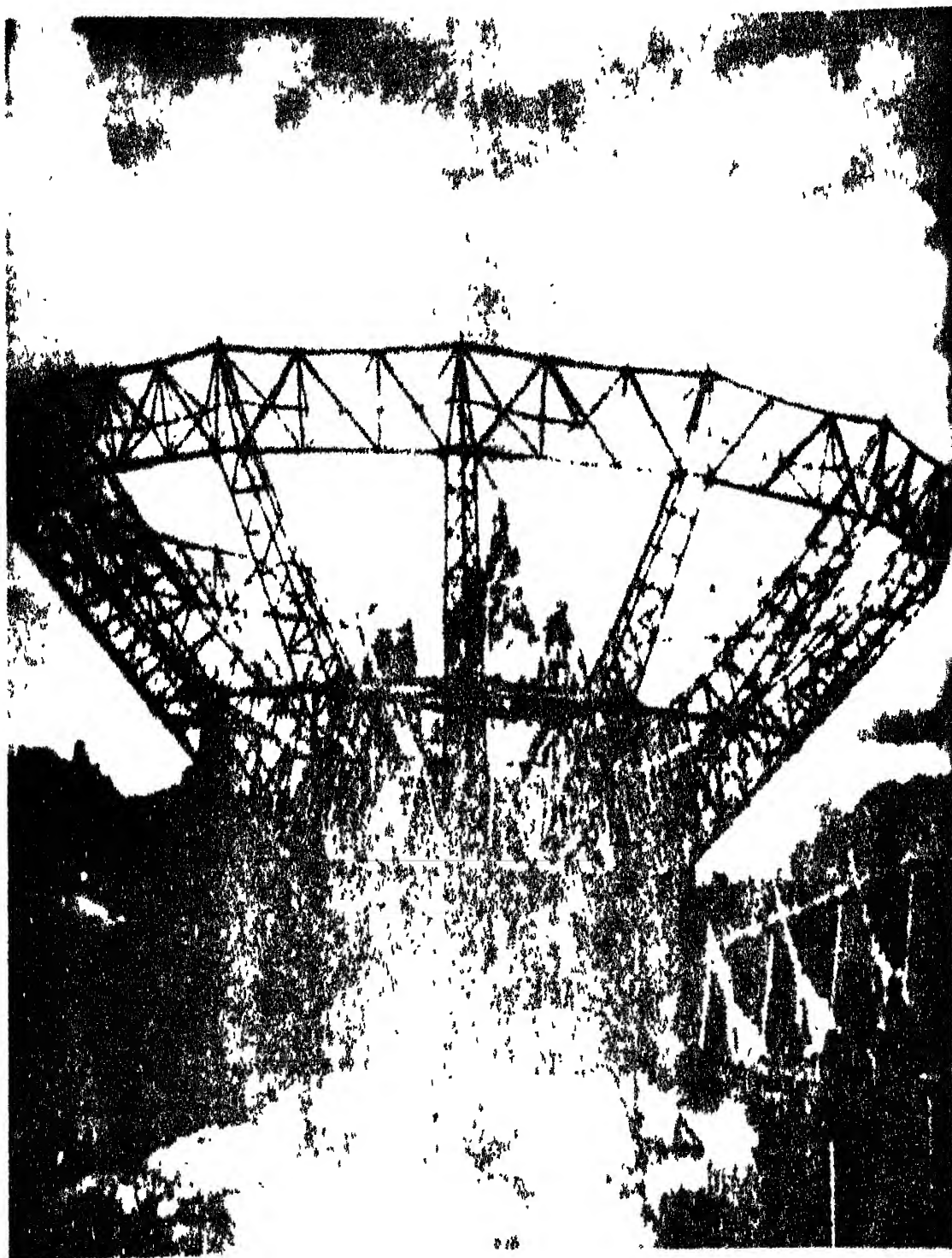


FIG. 20 (a) GMRT dish-note the novel design (see text) (b) distribution of GMRT dishes

radiation. Secondly, information of cosmological interest could emerge from extension of radio source counts and angular size statistics at meter wavelengths to flux levels of about 10 times lower than currently recorded and angular size-flux density relations some 100 times lower. Thirdly, a major (and perhaps the most significant aim) will be the highly redshifted 21cm line of clouds of primordial hydrogen with redshifts of about 3-10 from protoclusters or protogalaxies in early epochs of universe before galaxy formation.

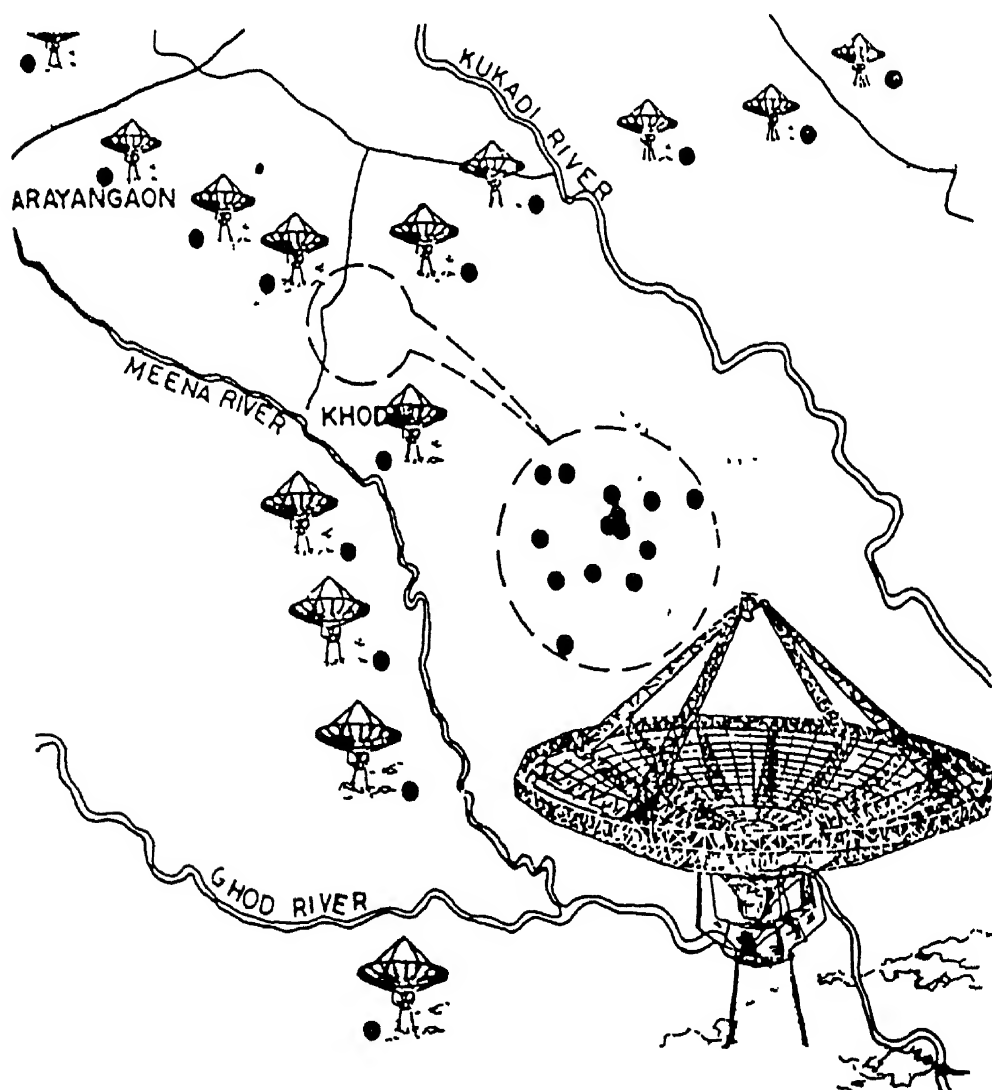


FIG. 20(b) Distribution of GMRT dishes.

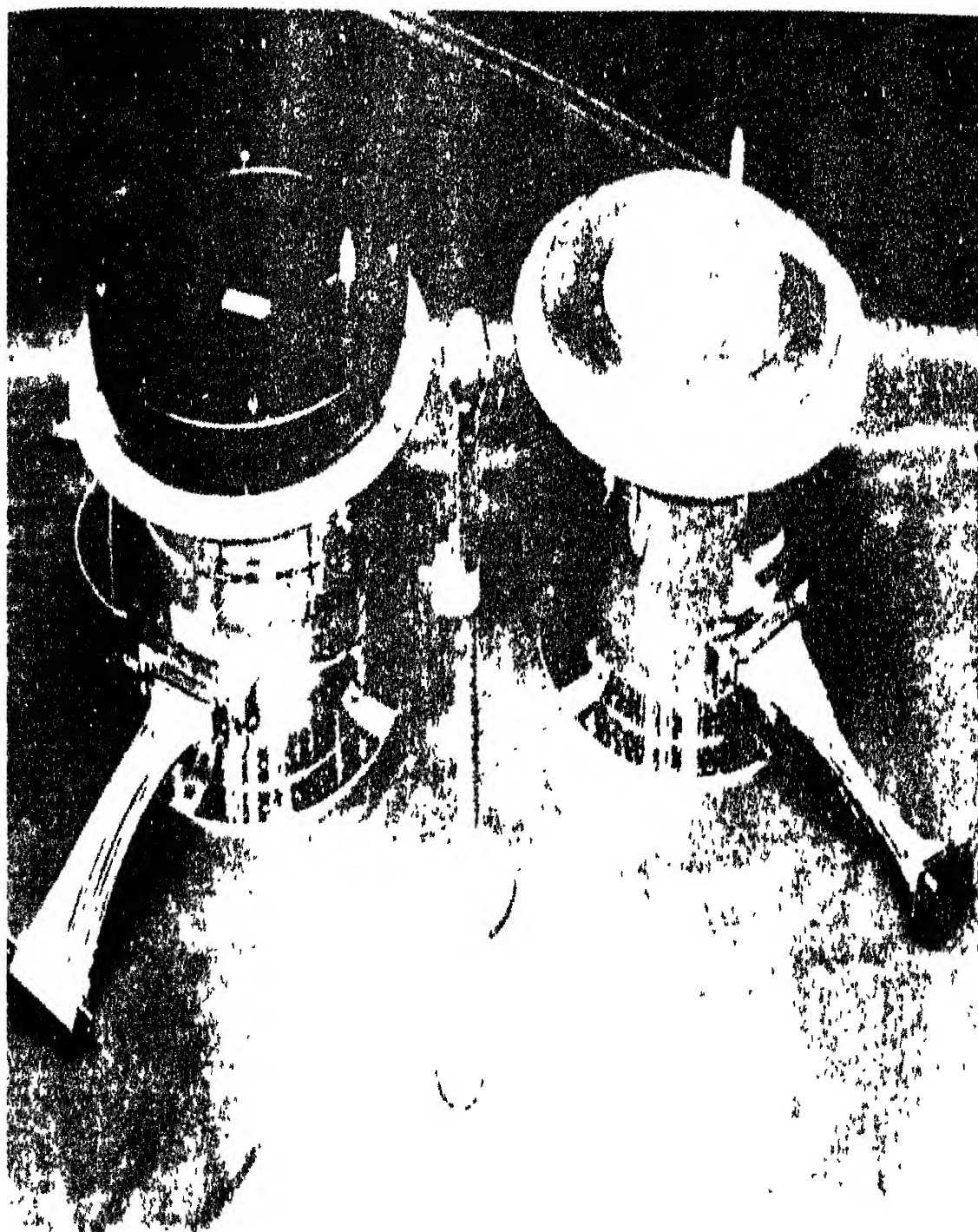


FIG 21 SROSS-C Retarding Potential Analyser Experiment (RPA) consisting of an iron sensor, electron sensor and potential probe)

### *SROSS-C Satellite*

The third major facility is the SROSS-C satellite recently launched. Although originally there was a plan of an independent aeronomy satellite, the finally accepted arrangement was one that included an ionospheric payload from NPL and a gamma ray burst experiment from ISRO. The satellite orbit has been selected (for the ionospheric experiment) to optimise information on upper atmosphere energetics and dynamics interactions. The orbit was originally scheduled to be  $450 \pm 50$  km; inclination of the orbit about  $46^\circ$ . The main objective is to look at the peak region of the low latitude ionosphere and to see how the neutral and ionization parameters are linked with sources that heat the region or produce motions. This means that one should not only measure the ionization levels but also the temperatures of the neutral particles ( $T_n$ ), of electrons ( $T_e$ ) and of ions ( $T_i$ ), and at the same time the nature and magnitude of electric fields and of energetic particles. The experiments thus have been designed to measure these parameters with retarding potential analyzers designed and built by NPL (Fig. 21). This is the first satellite ionospheric experiment mounted on an Indian satellite launched from India as a totally indigenous effort. For optimum utilization one should also have at the same time organized use of groundbased systems. In this case, support experiments include use of ionosondes, satellite radio beacon transmissions, Fabry-Perot interferometer, and use of one RH560 rocket flight (carrying RPAs to 50 km). Efforts are underway also to extend the range of the newly installed Tirupati ST radar to F region heights on a coherent mode.

### CONCLUDING REMARKS

We have given above a short review of progress in radio science in India. Clearly the scope, quality and perspectives have changed over the years. From the predominantly ionospheric programme of the early years, this has extended to cover radio meteorology and tropospheric propagation in the UHF, VHF and microwaves; radio astronomy with highly sophisticated systems in meter and decameter wavelengths on one side and

---

The satellite was launched on 20th May from SHAR with an initial orbit of 280 km (perigee) and 410 km (apogee). On June 17, the corresponding heights are 230 km and 370 km.



in mm wavelength on the other side and the exciting VLBI experiment for astronomy and geophysics; introduction of radio remote sensing with SAMIR in BHASKARA I and II and development of scatterometers on board aircrafts; introduction of new tools including meteor radars; VHF doppler backscatter radar and the exciting possibilities of the MST Radar; work on antennas and antenna system including Ooty synthetic telescope configuration and the conceptual innovativeness of GMRT; the beginning of signal processing, fibre optics and optical communication work, and a virtually explosive growth in semiconductor science and technology

New areas that are emerging which are yet to receive serious attention from Indian radio science communities include:-

- Interaction of electromagnetic waves with biological systems
- Radio geodesy: VLBI concept
- Electromagnetic Ecology  
(Changes in e.m. geophysical environment)
- Biophotons  
(ultraweak photon emission from biological tissues, cells, etc.)

One should also note that in the present concept radio frequencies are taken to extend to optical frequencies and so the role has widened considerably.

The dominant roles played by a few individuals and the impact of some major events at critical times (formation of the Radio Research Committee, the IGY and the IQSY, the Bhabha Committee Report, the IMAP, availability of new generation of satellites) are clear. To young radio scientists, the new national facilities now under design or construction offer unusual opportunities for frontier level work.

### ACKNOWLEDGEMENT

I would like to express my gratitude to Mrs Sudesh Mehra for typing of this article.

### References

1. S K Mitra, P Shyam, H Rakshit and B N Ghose *Nature* 131 (1933) 442
2. S K Mitra and H Rakshit *Phil Mag* 15 (1933) 20
3. S K Mitra and P Shyam *Nature* 135 (1935) 953

4. S K Chatterjee *Radio & Space Phys* **15** (1986) (5&6) 364
5. Amarjit Singh *Proc IRE (USA)* **43** (1955) 470
6. A P Mitra and C A Shain *J Atmos Terr Phys* **4** (1953) 2041
7. A K Saha, M Marabin and K K Mahajan *J Atmos Terr Phys* **25** (1963) 212
8. A P Mitra *Proc IGY Symp 1 CSIR (India)*1
9. Y V Somayajulu, T R Tyagi and V P Bhatnagar *Space Res* (1964); *Indian J Radio & Space Phys* **16** (1) (197) 136
10. Satyaprakash, S P Gupta and B H Subbaraya *Radio Sci* **4**(9) (1969) 791
11. Y V Somayajulu, S C Garg, R S Dabas, L Singh, T R Tyagi, B Lokanadhan, S Ramakrishna C Navneeth *Radio Sci* **19** (1984) 707
12. R G Rastogi, R P Sharma and V Shodar *Space Sci* **28** (1973) 713
13. S C Mazumdar, S K Sarkar and A P Mitra NPL (1977)
14. S K Sarkar, P K Pasricha, H N Dutta, B M Reddy and S M Kulsrestha NPL, New Delhi (1985)
15. R L Das *J Inst Telecomm Engr* **31** (1985) 126
16. S P Singal, J R Anand, B S Gera and S K Agarwal *CENTROP Rep.* No. 14, NPL, New Delhi (1974)
17. M K Rama *Ph.D Thesis Delhu Univ* (1978) M K Rama, G S Uppal and V R Ravindran *J Inst Electron Telecom Engrs* **25** (1979) 414-417
18. S C Mazumdar Ph.D. Thesis, Calcutta Univ (1974) S C Majumdar *Radio Electron Engg* **44** (1974) 63
19. G Swarup *Essays written in honour of Professor B C Sreekantan on his 60th Birthday* Ed R Cowsik Tata McGraw Hill, New Delhi
20. A Parmesh Rao and S Ananthakrishnan *Nature* **312** (1984) 707
21. S Ananthakrishnan, V K Kulkarni, D L Narayana, M R Sankararaman, A J Selvanayagam, H A Singal, N Subramanian, V Venkatanagarathinam, P Banerjee, A Sengupta, J E B Ponsonby and R W Procas *Indian J Rad Space Phys* **14** (1985) 17
22. S K Alurkar, R V Bhonsle and A K Sharma *Nature* **322** (1986) 439
23. Bhabha Committee on Electronics *Report of the Electronics Committee* Government of India February (1966)
24. J C Bose Collected Physical Papers Longman, Green & Co., New York, NY (1927)
25. S K Mitra, J N Bhar *Sci Cult* **10** (1) (1936) 788

26. S N Mitra *Proc Inst Electron Engg* **242** (1949) 441; *I J Radio & Space Phys* **15** (5 & 6) (1986) 295
27. H Rakshit *Phil Mag* **12** (1931) 897
28. G Swarup and V K Kapahi *Status Rep* June (1987) Radio Astronomy Centre Ooty
29. G Swarup, V K Kapahi, T Velusamy, S Ananthakrishnan, V Balasubramanian, Gopal Krishna, A Pramesh Rao, C R Subrahmanya and V K Kulkarni *Curr Sci* **60** (2) (1991) 79-94
30. G Swarup, S Ananthakrishnan, V K Kapahi, A P Rao, C R Subramanya and V K Kulkarni *Curr Sci* **60**(2) (1991) 95-105
31. S K Mitra *The Upper Atmosphere* The Asiatic Society, Calcutta



**Dinesh Mohan** (b. 14 July 1922) did C.E. (Hons.) (1943), Doc. Engng (h.c.) from University of Roorkee. He was INSA Senior Scientist (1985-88); Director, Central Building Research Institute (1962-82); Chairman, Bureau of Indian Standards, Civil Engineering Division

Dinesh Mohan developed under-reamed pile foundations in black cotton soil which provide trouble-free foundations in areas with expansive clays. Such foundations have been adopted in very large numbers for various types of structures all over India. Organized investigations on cost reduction of

primary school buildings and low-cost urban and rural housing. Based on these, he and his coworkers developed economical designs of such buildings which have been put up in large numbers in various parts of the country.

Dinesh Mohan is Fellow, Institution of Engineers (India) and Indian National Academy of Engineering; Life Member, American Society of Civil Engineers, President, Indian Geotechnical Society (1961-71); Vice-President, International Society of Soil Mechanics and Foundation Engineering (1976-80) and International Council of Building Research Studies and Documentation (1974-77), Honorary Member, International Council of Building Research & Documentation (CIB) (1988). He is the recipient of UGC National Lecturer (1976-77), Shanti Swarup Bhatnagar Medal (INSA) (1988).

*Dinesh Mohan was elected to the fellowship of the Academy in 1970*

**Address:** RI/19, Rajnagar, Ghaziabad 201 002; Tel: 8713838.

# **BLACK COTTON SOILS—HIGHLY EXPANSIVE CLAYS OF INDIA**

DINESH MOHAN FNA

## **INTRODUCTION**

India has large tracts of expansive soils known as black cotton soils. This name derives from their black colour and cotton being grown in many regions on these soils. The major area of their occurrence is Central India, South of Vindhya range, covering an area of about 0.8 million sq. km., thus forming about 20 percent of the total land area of the country.

The Indian black cotton soils are generally heavy clays containing predominance of clay mineral montmorillonite, exhibiting high shrinkage and swelling characteristic. During shrinkage, due to drying, there is a formation of hexagonal columnar structure with vertical cracks upto 8cm wide at the ground level, extending upto 2.5m depth. Shrinkage in horizontal direction is nearly two third of the total volumetric shrinkage. Similar soils occur in other countries also e.g. 'Chernozems' of U.S.S.R., 'Badole' of Japan, 'Pampus of Argentina', 'Tirs' of Morocco, 'Margilatic' of Indonesia and black earths of Australia, Java & Sumatra. Vast area near Irbid (Jordan) are also full of expansive clays. Geologically their formation is associated with basalts but their occurrence on granite, shale, sandstone and slate is also recognised. They occur both as residual and transported. In the former case, where the parent rock lies underneath, their depth is shallow, averaging about 1m, but in low lying and flat areas when they develop over alluvium, after being transported by the wind and water, they go deep and usually average 5m.

## **BASIC SOIL PROPERTIES & THEIR CORRELATION WITH ENGINEERING PROPERTIES**

The author procured samples of black cotton soil from twenty different places in India (Fig. 1) covering practically all the regions where such soils exist. The depth of the samples varied from 1-3m. The liquid limit ranged between 46 and 97, plasticity index between 22 and 49 and shrinkage limit between 11 and 14. The plot on the plasticity chart (Fig. 2) showed that practically all the soils fall above the Casagrande 'A' line.

The organic content varied from 0.4 to 2.4 per cent and specific gravity had an average value of 2.7.

A straight line relationship was observed between shear strength (plotted to log scale) and liquidity index (Fig. 3). For this experiment the soil was compacted at the plastic limit in a Proctors mould by the standard Proctor rammer. Cylindrical test specimen of diameter 5cm and height 10cm were obtained from the compacted soil mass in the mould by pushing in a thin walled steel sampling tube. These were subjected to gradual air drying and compressive strength was determined at four

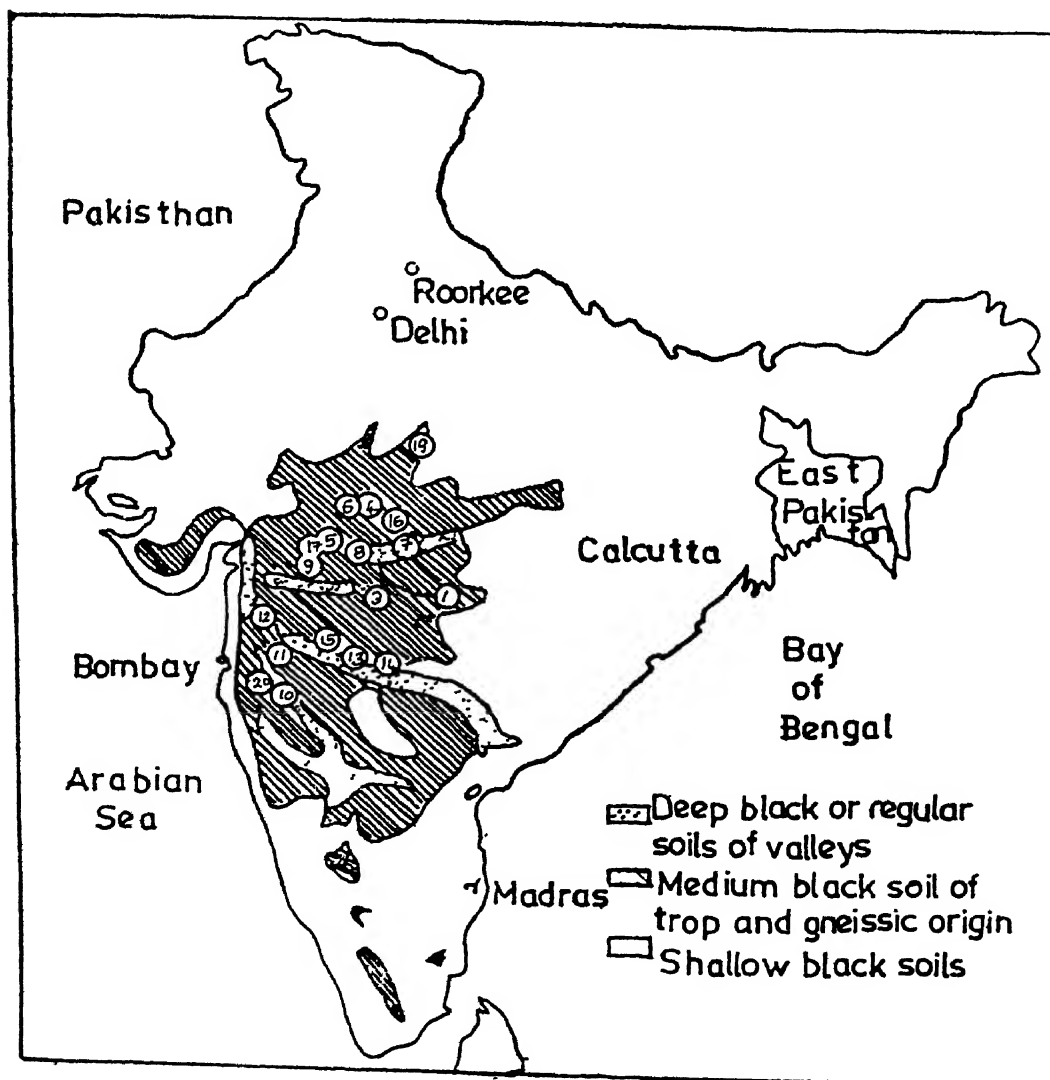


FIG. 1 Soil map of India showing location of black soils

different moisture conditions in a decreasing order in a compression testing machine working at a uniform rate of loading. The curve has been plotted only upto liquidity index range of 0.3 which gives the minimum moisture content at which black cotton soils usually exist at a depth of 1m. The liquidity index has a negative value in all cases as the moisture content was below the plastic limit. The concentration of points is along a straight line and at the plastic limit the strength of the soil is very low.

No appreciable difference was found in strength between undisturbed and remoulded samples (Table 1). This was checked by taking a portable unconfined compression test equipment to the site and testing undisturbed and remoulded samples from various depths. The sensitivity of Indian black cotton soils is therefore close to unity.

Table I

| Location   | Depth<br>(ft.) | Shear strength<br>undistributed(psi) | Shear strength<br>remoulded(psi) |
|------------|----------------|--------------------------------------|----------------------------------|
| Powerkhera | 4              | 19                                   | 21                               |
|            | 8              | 21                                   | 17                               |
|            | 12             | 20                                   | 18                               |
| Indore     | 4              | 17                                   | 16                               |
|            | 8              | 22                                   | 23                               |
|            | 12             | 29                                   | 29                               |

Differential free swell test was used to get a comparative idea of the swelling potential of the soil. The normal free swell test suggested by Holtz and Gibbs consists of gently pouring 10 cc of oven dried soil, passing 36 BS sieve (Particle size less than 0.42mm), into distilled water and noting the increase in volume of the soil. The main drawback of the method is lack of uniformity in packing and long time required for the soil to come to constant volume in a soaked state. The differential free swell test developed by RDSO<sup>8</sup> overcomes this drawback. In this method, 10gm of two oven-dried samples of soil are soaked, one in distilled water and the other in kerosene oil or some other non-polar liquid, their respective volumes being measured. The difference in volume expressed as a percentage of the volume in kerosene oil gives the differential free swell. The values were found to range from 28 to 122. A curve was plotted between the differential free swell on one hand and volume expansion and

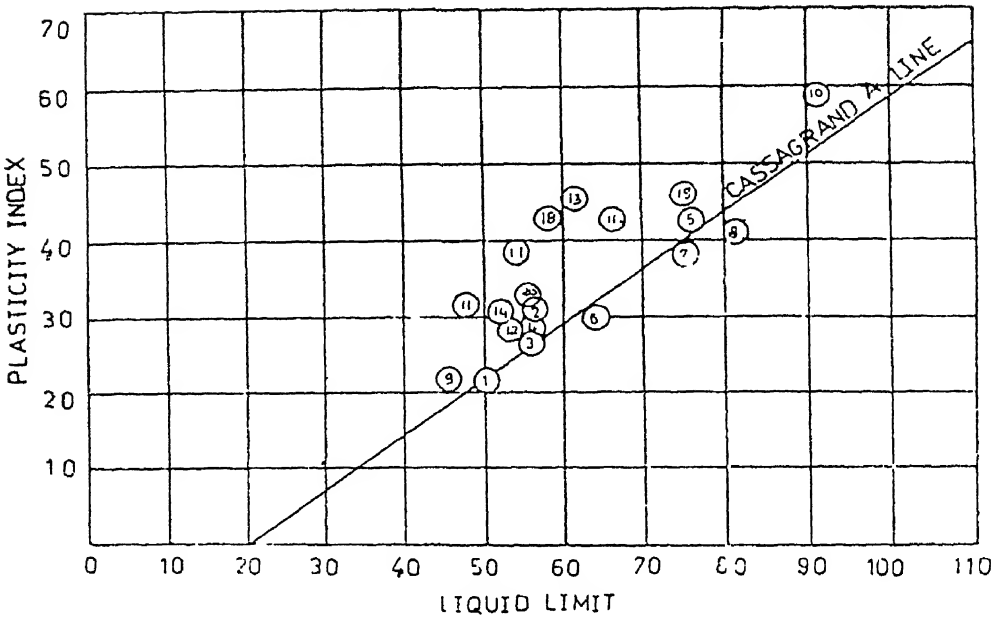


FIG. 2 Relation between liquid limit and plasticity index of black cotton soil samples

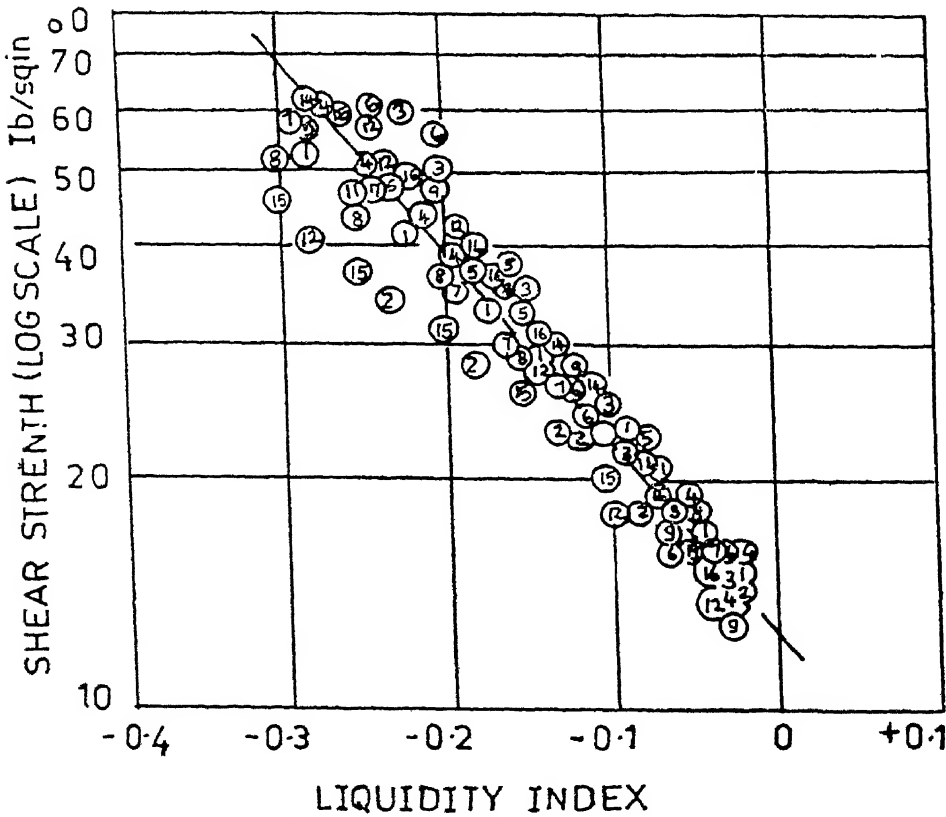


FIG. 3 Shear strength versus liquidity index



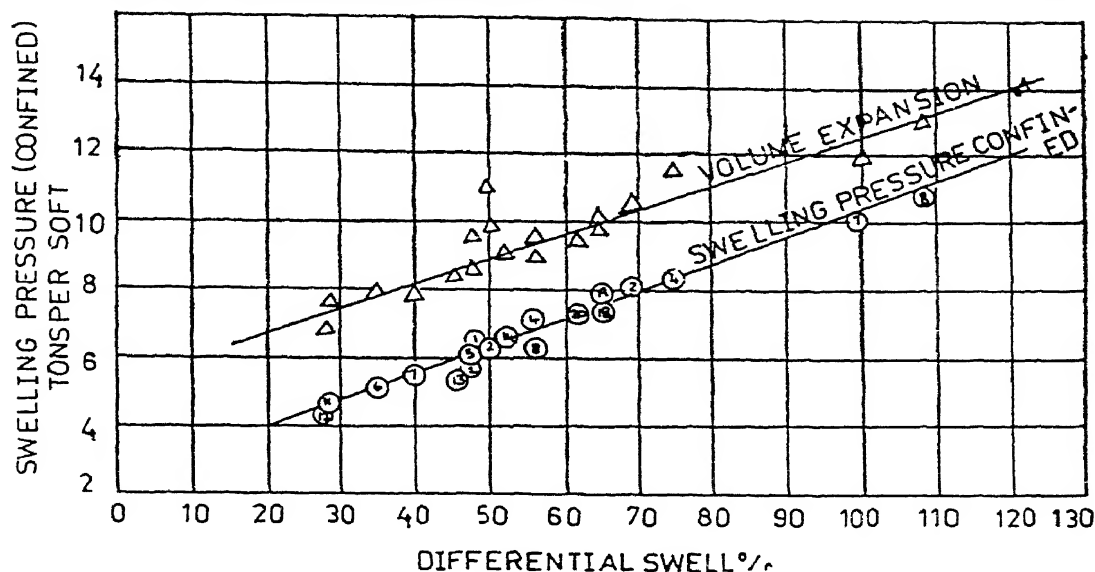


FIG. 4 Relation between differential swell, and swelling pressure and volume expansion of black cotton soil samples

swelling pressure on the other. A straight line relationship was observed (Fig. 4).

### FOUNDATION DESIGN AND CONSTRUCTION

Foundation of buildings and other structures on Indian Bank cotton soils have been a matter of great concern to the engineers and builders. It is heart-rending to see some of the structures cracking badly within a year after their construction.

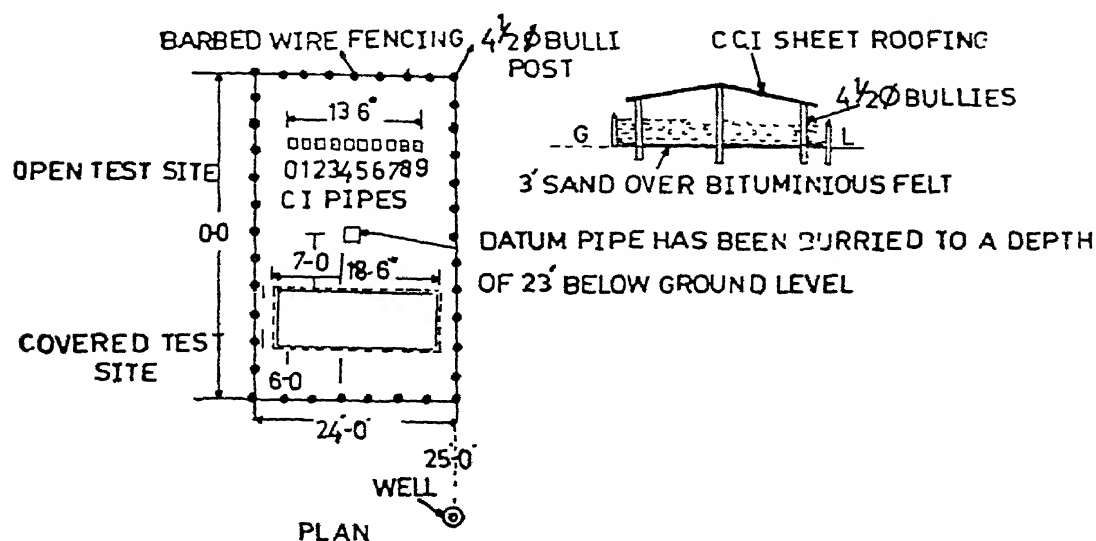


FIG. 5 Plan of the test site

Until 1955, the normal methods for constructions of foundations on black soils were -

- (1) Provision of reinforced concrete bands at plinth and lintel level.
- (2) Provision of reinforced concrete raft or beam to support the superstructure.
- (3) Removing the black soil entirely or to a considerable depth and backfilling the trench with cohesionless soils.

Method (1) was not very effective and method (2) was costly. Method (3) is also uneconomical where depth of black soil exceeds 1m.

The main reason leading to the failure of foundations in black soils is the differential movement of the structure with uneven ground movements due to alternate swelling and shrinkage of the soil. The vertical ground movements decrease with depth and are negligible at a

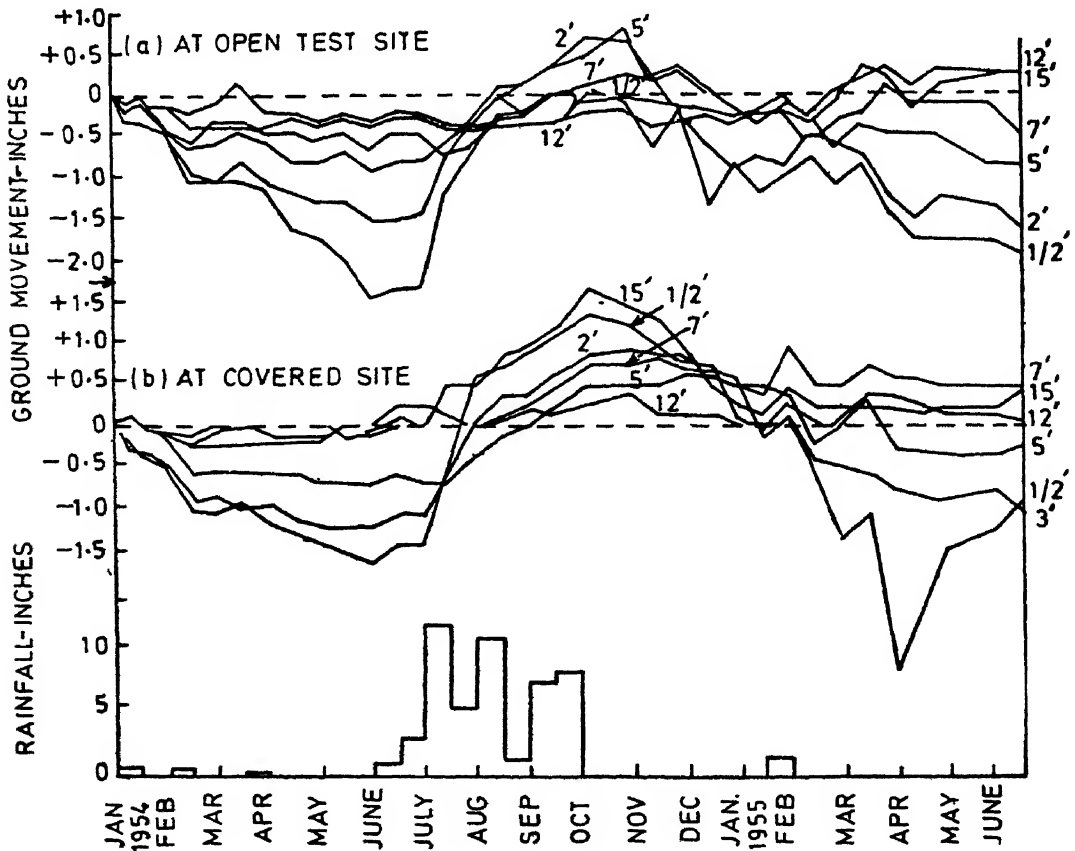


FIG 6 Record of vertical ground movement at Hoshangabad

particular depth. Investigations were carried out by CBRI to locate this depth in two different regions with deep layers of black soils. Test sites were established at both these places and each of the sites had a row of G.I. pipes anchored at different depths from 6 in (15cm) to 15 ft. (4.5m) deep. There were two such rows at each site, one in the open and the other under a shed and protected by bituminous felt to simulate the cover produced by a building (Fig. 5). Fortnightly levels were recorded of each pipe over a period of 13 to 18 month and these have been plotted in Fig.6 and 7. From a study of the curves it would be seen that 12 feet (3.5m) is the depth at which the vertical ground movements are inappreciable.

Another test site 25mX35m was set up after a lapse of about 15 years in a black cotton soil area where the top 2.7m was black silty clay overlying yellow clay which extended to 5m and beyond<sup>11</sup>. A number of surface movement indicators, depth gauges and *in situ* swelling pressure measuring instruments were installed and measurement were carried out over a period of 3 years (Fig. 8). Maximum value of ground movement at the surface was found to be 65mm which decayed to a negligible value at about 5m depth. Bulk of the movement was found to occur within the top 2m and 10 per cent of maximum occurred at nearly 3/4 of the depth of negligible heave. A depth of  $5 \times 3/4 = 3.75\text{m}$  could therefore be taken as a safe depth for foundations. *In situ* swelling pressure studies were carried out at depths varying from 0.5m to 2.5m using individual beam set up (Fig. 9). In order to accelerate swelling, the area was continuously flooded for a period of five months. Out of a total six swelling pressure setups two measured *in situ* swelling pressures by accelerated tests. Swelling pressure values as high as 5.5 Kg/cm<sup>2</sup> was observed with depth (D) which could be expressed by the following relationship  $6_s = 8 - 1.6D$ . Having found 3.50-3.75m as depth of inappreciable ground movement, a design was developed of under-reamed pile foundations with the belled out portion anchored at a depth of 3.5m or earlier if water table or a stable strata was encountered. The boring for the pile was carried out by a spiral auger and its bottom was under-reamed to about  $2\frac{1}{2}$  times the shaft diameter. For this purpose a portable, hand-operated under-reaming tool was designed (Fig. 10). The tool is simple to fabricate in a workshop. It consists of an assembly of four steel blades fixed around a central shaft and a bucket to hold the cut soil. The blades widen out as the shaft is pressed downwards and by rotating them the bore hole is widened. The shaft has a number of holes and a pin<sup>12</sup> inserted in a particular hole controls the maximum diameter of the bell.

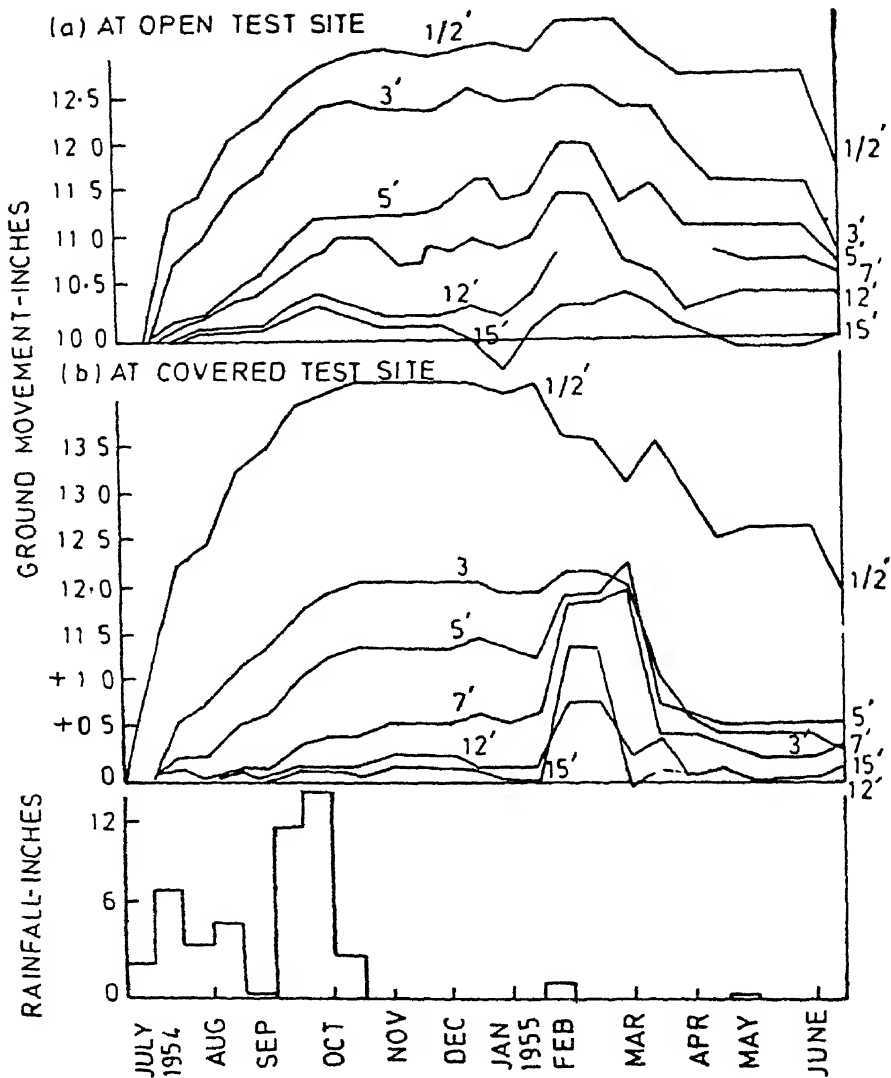


FIG 7 Record of vertical ground movement at Indore

The pile diameter varies from 20 to 50cm depending on the load it is expected to carry. For small works, in out of the way places, manual operation is usually preferred but for larger diameter and deeper piles, a mechanised boring rig is normally necessary.

The piles are provided at appropriate locations keeping in view the layout of the building and the load carrying capacity of the piles. As far as possible all piles are uniformly loaded and the spacing so adjusted as to keep the door and window openings mid-way between two piles. A typical layout of a residential building and details of piles is given in Fig. 11(a) and a section through the pile foundation is given in Fig. 11(b). It would

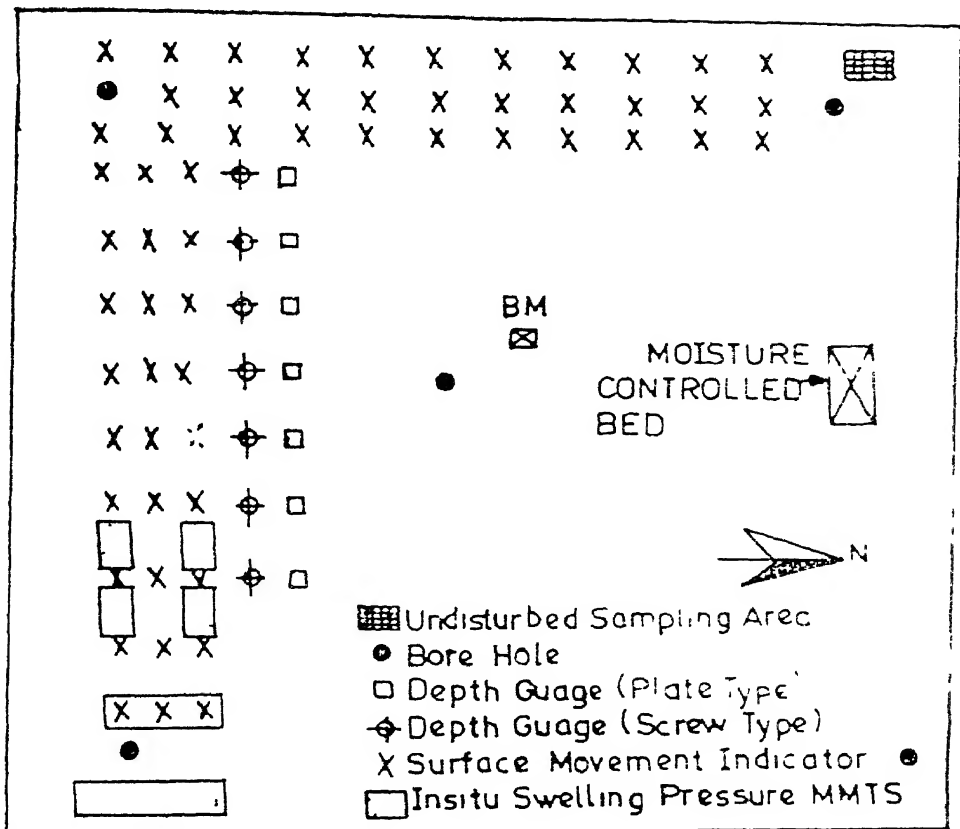


FIG 8 Site for field experiment showing layout of measuring instruments

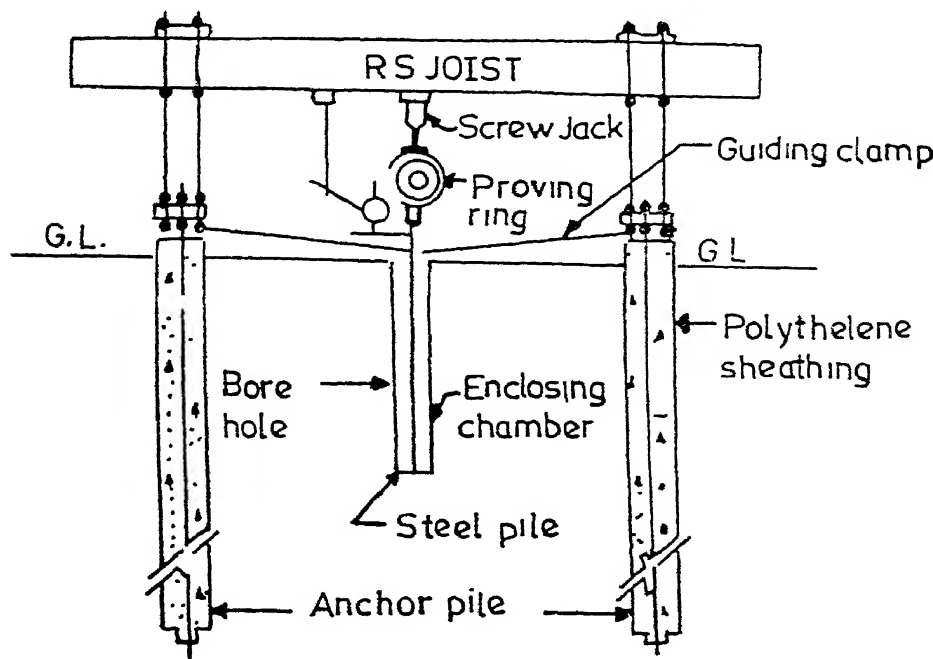


FIG 9 Swelling pressure measurement using individual beam set up

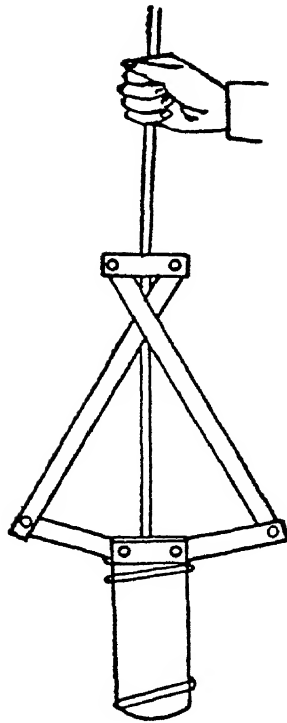


FIG 10 Under-reaming tool

be seen that the grade beam is kept 3in (8cm) clear of the ground so that the soil does not heave against it. These beams, carrying the masonry superstructure are designed for a bending moment of  $WL/50$  to allow for panel action in the masonry. The shuttering supporting the beams is therefore not removed for about a week.

The bearing capacity ( $Q_u$ ) of under-reamed piles can be determined by the following expression normally used for bored piles,

$$Q_u = A_p N C_p + \alpha C A_s (l),$$

- where,
- $A_p$  = area of the pile tip (under-reamed base of the pile)
  - $N$  = bearing capacity factor (may be taken 9)
  - $C_p$  = undisturbed shearing strength of the soil at the bearing level
  - $\alpha$  = reduction factor for bored piles
  - $C$  = average undistributed shearing strength of the soil along the pile length
  - $A_s$  = Surface area of the pile shaft

The value of reduction factor ( $\alpha$ ) may be taken as 0.5. This has been determined by carrying out a series of loading and pull out tests on cast-in-situ bored concrete piles in black cotton soils at four different sites in India.<sup>10</sup>

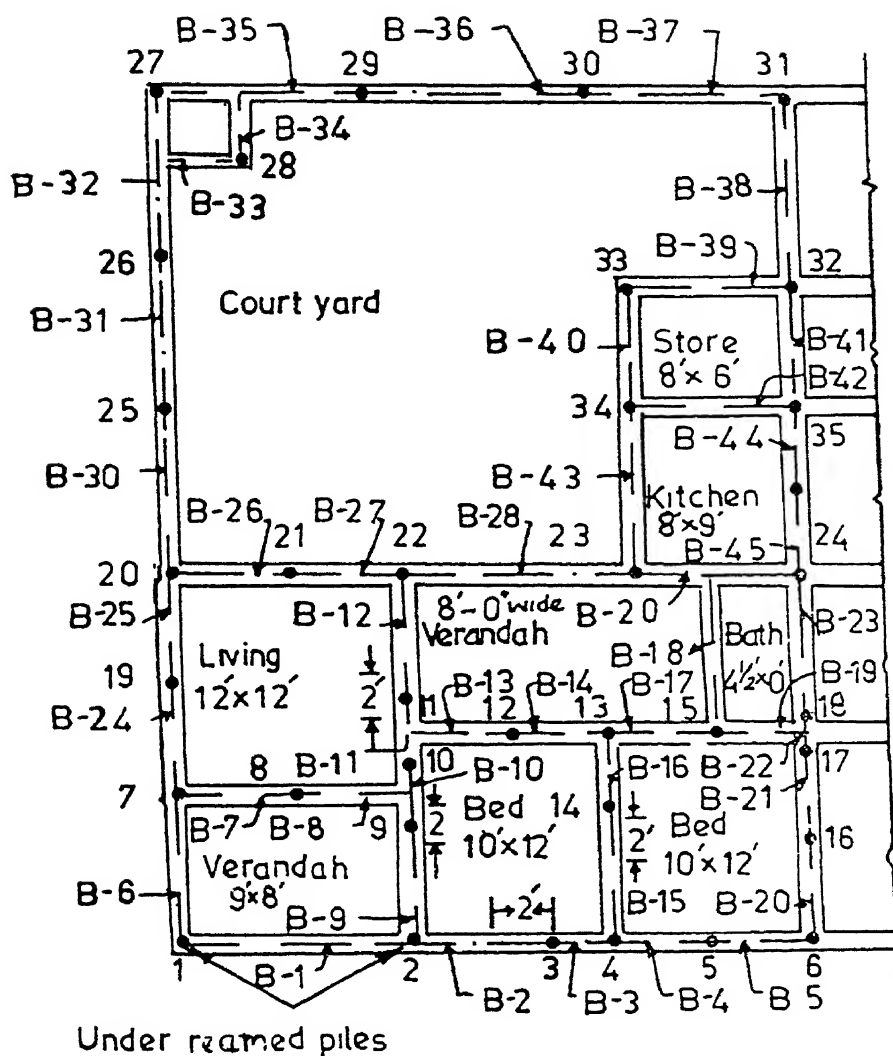


FIG 11 (a)

The adoption of under-reamed pile in expansive soils has resulted in economy to the extent of 30-60 per cent when compared to traditional strip footings. An additional advantage is that the process is quick and no extra excavation or backfilling is required. This provides better and more uniform conditions for floor finishes adjacent to the walls.

A full scale load test 10 was carried out on a 7.6m long and 45cm stem diameter multi-under-reamed pile, having four bulbs, each 112.5cm diameter, in a deep layer of black cotton soils, having undrained cohesion varying from 0.9 to 1.45kg/cm<sup>2</sup>, liquid limit from 65 to 70 and plasticity

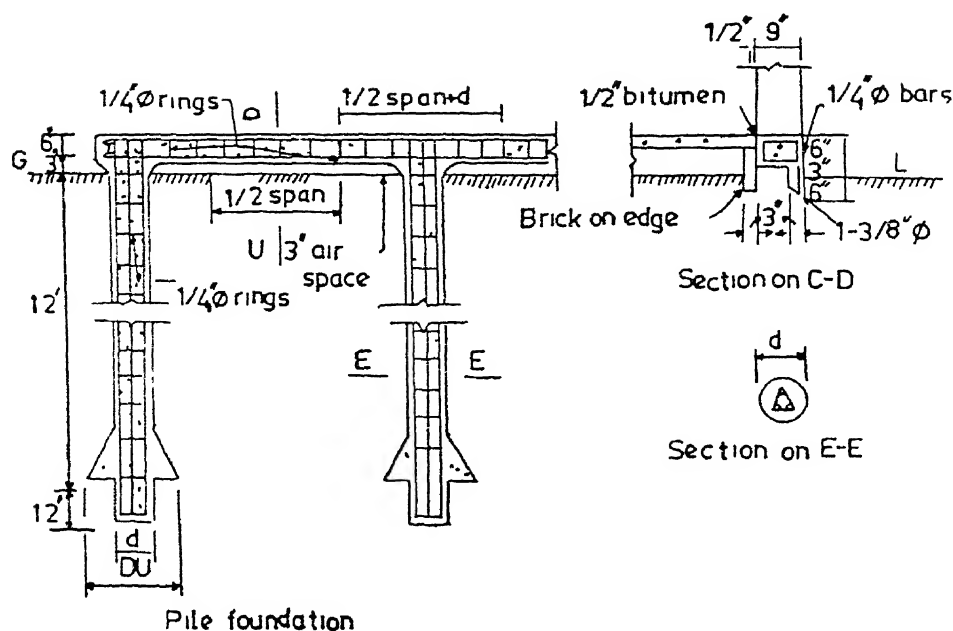


FIG 11 (b)

index from 35 to 50. Twelve under-reamed piles, arranged in a circle, were used as anchors for the loading frame designed to carry a load of 400 tonnes. A sectional view of the test set up is shown in Fig 12. Load carrying capacity of the test pile as worked out from table *Appendix-A* was found to be 290 tonnes. The computed value, using bearing capacity formula (1), gave a value of 328 tonnes. The pile could only be tested upto 300 tonnes at which the settlement was only 14 mm. Further testing was not possible due to yielding of the anchor piles. It is clear from the load test that the ultimate capacity of the pile is more than 300 tonnes and could be close to the computed value of 328 tonnes.

### CONCLUDING REMARKS

The use of under-reamed piles in India is now an accepted practice in black cotton soil areas. It is no longer restricted to expansive soils or buildings alone but is used in other types soil conditions such as filled up grounds and poor soils overlying firm strata, foundations subjected to uplifts and thrusts such as transmission line towers, underground tanks etc. A private developer has even adopted them for normal grounds in view of their speedy construction. An Indian Standard code of practice [IS : 2911 (Part III) 1980] on under-reamed piles has been brought out to help the



designer and a safe load table provided in the code has been given in Appendix-A

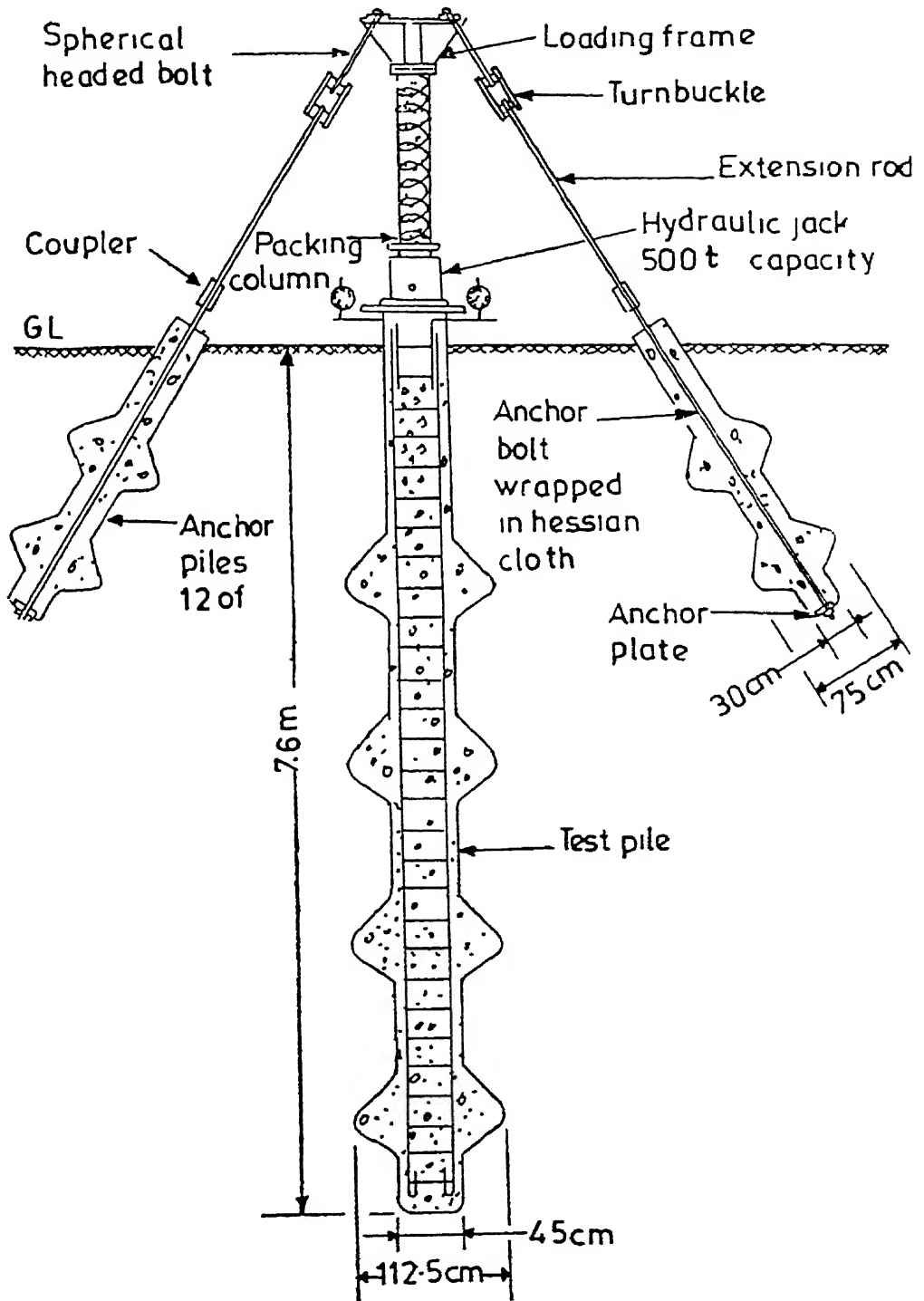


FIG 12 Sectional view of the pile load test set-up

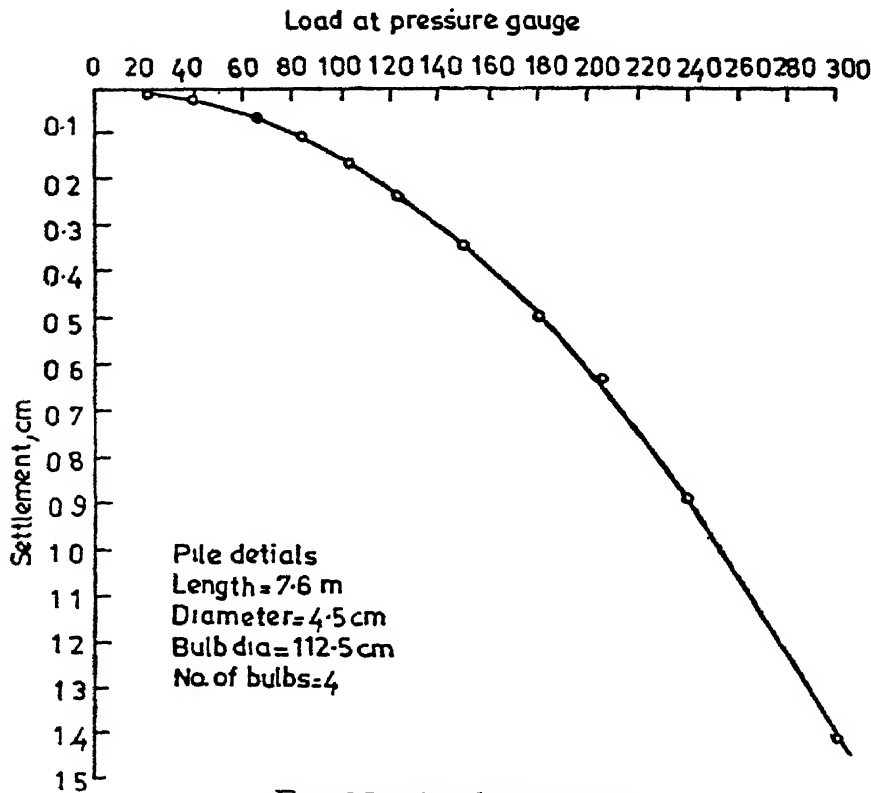


FIG 13 Load settlement curve

It can therefore be safely claimed that under-reamed pile foundations have provided a foolproof, economical and quick to construct solution for foundation in expansive clays in India.

## References

1. Ray Chaudhri M Sulaiman and A B Bhuiyn Physico-chemical and minerological studies of black and red soil profile near Coimbatore *Indian J agric Sci* 13(1943)
2. F L D Wooltorton *The Scientific Basis of Road Design* Edward Arnold Ltd London (1954)
3. K J Vogel Foundations in shrinkable soils *Ministry of Railways Govt of India Tech paper No. 329* (1954)
4. N Shanmugam Building foundations in black cotton soil *J Indian natn Soc SMFE* 1(3) (1962) 60
5. D Mohan Consolidation and characteristics of Indian black cotton soils *Proc int Conf SMFE London* (1957) 74-76
6. Dinesh Mohan and R K Goel Swelling pressure and volume expansions of Indian black cotton soils *J Inst Engrs (India)* XL(2) (1959) 58-62
7. W G Holtz and H J Gibbs Engineering properties of expansive clays *Trans ASCE Paper No 2814* (1956) p 641

8. Anon Some classification tests for shrinkage and swelling soils *Indian Railway Tech Bull* 15(128) (1958) p 31
9. Dinesh Mohan and Giri Raj Singh Jain Under-reamed pile foundation on black cotton soils *Indian Conf News* Dec 25-28 (1955)
10. Dinesh Mohan and Subhas Chandra Frictional resistance of bored piles in expansive clays *Geotech* (1961) 294-301
11. S P Gupta and R K Bhandari *In-situ* engineering behaviour of an Indian black cotton soil deposit *Proc VII Regional Conf ISSMFE Haifa (Israel)* (1983) 19-23
12. Dinesh Mohan and G S Jain Pile testing using raked under-reamed pile *Gr Engng* (1979) April 47-52

## Appendix A

*Extract from Indian Standard Code of Practice for Under-reamed Piles-IS 2911 (Part III)-1980*

TABLE

*Safe load for vertical bored cast in situ under-reamed piles in sandy and clayey soils including black cotton soils*

| Size              |                        | Length              |                     | Mild Steel Reinforcement    |                                 | Compression              |                     | Safe Loads in Uplift Resistance |                           |                          |                     |                            |                               | Lateral Thrust      |              |      |
|-------------------|------------------------|---------------------|---------------------|-----------------------------|---------------------------------|--------------------------|---------------------|---------------------------------|---------------------------|--------------------------|---------------------|----------------------------|-------------------------------|---------------------|--------------|------|
| Dia meter of pile | Under reamed dia meter | Single under-reamed | Double under-reamed | Longitudinal reinforce-ment | Rings spacing of 6 mm dia rings | Single under-ream-<br>ed | Double under-reamed | In-crease per 30cm length       | De-crease per 30cm length | Single under-ream-<br>ed | Double under-reamed | In-crease per 30 cm length | De-crease reamed 30 cm length | Single under-reamed | Double under |      |
| cm                | cm                     | m                   | m                   | No                          | Dia mm                          | cm                       | t                   | t                               | t                         | t                        | t                   | t                          | t                             | t                   | t            |      |
| (1)               | (2)                    | (3)                 | (4)                 | (5)                         | (6)                             | (7)                      | (8)                 | (9)                             | (10)                      | (11)                     | (12)                | (13)                       | (14)                          | (15)                | (16)         | (17) |
| 20                | 50                     | 3.5                 | 3.5                 | 3                           | 10                              | 18                       | 8                   | 12                              | 0.9                       | 0.7                      | 4                   | 6                          | 0.65                          | 0.55                | 1.0          | 1.2  |
| 25                | 62.5                   | 3.5                 | 3.5                 | 4                           | 10                              | 22                       | 12                  | 18                              | 1.15                      | 0.9                      | 6                   | 9                          | 0.85                          | 0.70                | 1.5          | 1.8  |
| 30                | 75                     | 3.5                 | 3.5                 | 4                           | 12                              | 25                       | 16                  | 24                              | 1.4                       | 1.1                      | 8                   | 12                         | 1.05                          | 0.85                | 2.0          | 2.4  |
| 37.5              | 94                     | 3.5                 | 3.75                | 5                           | 12                              | 30                       | 24                  | 36                              | 1.8                       | 1.4                      | 12                  | 18                         | 1.35                          | 1.10                | 3.0          | 3.4  |
| 40                | 100                    | 3.5                 | 4.0                 | 6                           | 12                              | 30                       | 28                  | 42                              | 1.9                       | 1.5                      | 14                  | 21                         | 1.45                          | 1.15                | 3.4          | 4.0  |
| 45                | 112.5                  | 3.5                 | 4.5                 | 7                           | 12                              | 30                       | 35                  | 52.5                            | 2.15                      | 1.7                      | 17.5                | 25.75                      | 1.60                          | 1.30                | 4.0          | 4.8  |
| 50                | 125                    | 3.5                 | 5.0                 | 9                           | 12                              | 30                       | 42                  | 42                              | 63                        | 1.9                      | 21                  | 31.5                       | 1.80                          | 1.45                | 4.5          | 5.4  |

## APPENDIX

1. The safe bearing, uplift and lateral loads for under reamed piles given in the Table on p. 28 apply to both medium compact ( $10 < N < 30$ ) sandy soils and clayey soils of medium ( $4 < N < 8$ ) consistency including expansive soils. The values are for piles with bulb diameter equal to two-and-a-half times the shaft diameter.

The columns (3) and (4) of Table provide the minimum pile lengths for single and double under-reamed piles, respectively, in deep deposit of expansive soils. Also the length given for 375mm diameter double under-reamed piles and more in other soils are minimum. The values given for double under-reamed piles in columns (9) and (13) are only applicable in expansive soils. The reinforcement shown is mild steel and it is adequate for loads in compression and lateral thrusts (Columns (8), (9), (16) and (17)). For up lift (Columns (12) and (13)), requisite amount of steel should be provided. In expansive soils, the reinforcement shown in Table is adequate to take upward drag due to heaving up of the soil. The concrete considered is M 15.

2. Safe loads of piles of lengths different from those shown in Table can be obtained considering the decrease or increase as from Columns 10, 11, 14 and 15 for the specific case
3. Safe loads for piles with more than two bulbs in expansive soils and more than one bulb in all other soils (including non-expansive clayey soils) can be worked out from Table by adding 50 per cent of the loads shown in columns (8) or (12) for each additional bulb to the values given in these columns. The additional capacity for increased length required to accommodate bulbs should be obtained from columns (10 and (14)
4. Values given in columns (16) and (17) for lateral thrusts shall not be increased or decreased for change in pile lengths. Also for multi-under-reamed piles the values shall not increase than those given in column (17). For longer and /or multi-under reamed piles higher lateral thrusts may be adopted after establishing from field load tests.
5. For dense sandy ( $N > 30$ ) and stiff clayey ( $N > 8$ ) soils, the safe loads in compression and uplift obtained from Table may be increased by 25 per cent. The lateral thrust values should not be increased unless the stability and strength of top soil (strata upto a depth of about three times the pile shaft diameter) is ascertained and found adequate. For piles in loose ( $4 < N < 10$ ) sandy and soft ( $2 < N < 4$ ) clayey soils, the safe loads should be taken as 0.75 times the values shown in the Table. For every loose ( $N < 4$ ) sandy and very soft ( $N < 2$ ) clayey soils the values given in the Table should be reduced by 50 per cent.
6. The safe loads obtained from Table, should be reduced by 25 per cent if the pile bore holes are full of subsoil water or drilling mud during concreting.
7. The safe loads in uplift and compression given in Table or obtained in accordance with 2 to 6 should be reduced by 15 per cent for piles with bulb of twice the stem diameter. But not such reduction is required for lateral loads shown in Table.

The safe loads in Table and the recommendations made to obtain safe load in different cases (2 to 8) are based on extensive pile load tests obtained may be taken equal to two-thirds the loads corresponding to deflection of 12mm for loads in compression and uplift. The deflections corresponding to respective safe loads will be about 6mm and 4mm. The deflection at safe lateral load will be about 4mm. The values given in Table will be normally on conservative side. For working out ultimate compressive and uplift loads, if defined as loads corresponding to 25mm deflection on load-deflection curve, the value obtained from Table can be doubled. In case of lateral thrust twice the values in Table should be considered corresponding to deflection of 12 mm only.

9. For piles/subjected to external moments and/or larger lateral loads than those given in Table the pile should be designed properly and requisite amount of steel should be provided.



**Bimal Kumar Bachhawat** (b. 16 August 1925) did Ph.D. (1953) from University of Illinois, USA. He is Bhatnagar Fellow (1990- ) and was formerly Dean, (1988-90) at Faculty of Interdisciplinary and Applied Sciences, University of Delhi, South Campus, New Delhi.

Bachhawat's discovery of the biosynthetic pathway of cerebroside-3-sulphate led to the understanding of the mechanism of the synthesis of this complex glycolipid. The original finding that the aryl-sulphatase A is the deficient enzyme in the in-born error of metabolism (metachromatic leucodystrophy) led to the development of a simple diagnostic method for this disease. His discovery (with M.J. Coon) of the enzyme HM6COA lyase led to the basic understanding of the ketone body synthesis in animals. The discovery of a new enzyme CMP-n, acetylneuraminic acid degrading enzyme, shed light on its regulatory role in the biosynthesis of membrane sialic acid. The receptor-ligand interaction with the liposome as a model membrane led to the understanding of the density of receptor on the cell surface and its role in ligand interaction. Recently, Bachhawat and his group have developed a liposomal formulation of Amphotericin-B which can effectively cure systemic aspergillosis in experimental animals. He has Authored more than 20 books/chapters.

Bachhawat is Fellow of Indian Academy of Sciences and National Academy of Sciences (India); He was Member, Council (1975-77) and Vice-President (1987-88), INSA. He is the recipient of Shanti Swarup Bhatnagar Prize (1962); Amrut Mody Research Award (1974); Golden Jubilee Gold Medal (Institute of Science) (1976); J.C. Bose Award (1980); Bashambar Nath Chopra Memorial Lecture Award (INSA) (1977); FICCI Award (1982); B.C. Guha Memorial Lecture Award (INSA) (1984); Birla Smarak Khosh (1986); Shanti Swarup Bhatnagar Medal (INSA) (1991); Padma Bhushan (1990).

*Bimal Kumar Bachhawat was elected to the fellowship of the Academy in 1973.*

**Address:** Bhatnagar Fellow, Department of Biochemistry, University of Delhi, South Campus, New Delhi 110 021; Tel: (011) 678876, 6882231 (Off); 606207 (Res); Fax: 91-11- 6885270.

# ENHANCED STABILITY OF HYDROPHILIC POLYMER CONJUGATED ENZYMES, BOTH IN CIRCULATION AND INTRACELLULARLY : AN APPROACH TO ENZYME THERAPY

GARGI MITRA, SOFIA MUMTAZ and BIMAL K BACHHAWAT

*Intravenous injection of enzymes and other agents may lead to their prompt removal from the circulation due to hepatic uptake or rapid degradation by proteolytic enzymes. Conjugation with dextran has been shown to endow these molecules with a capacity to persist in the circulation as well as inside the cell for extended periods of time. Presumably, the altered topology and physicochemical properties of the conjugated proteins are responsible for such effects which render them extremely useful from the therapeutic point of view.*

## INTRODUCTION

The past two decades have witnessed much attention being directed to the covalent modification of bioactive molecules for imparting *in vivo* stability to the constructed conjugates. The disposition and pharmacokinetics of a molecule conjugated with dextran would be expected to be dictated predominantly by the dextran carrier. Dextran, a polymer of  $\alpha$ -D-glucose units ( $\alpha 1 \rightarrow 6$ ) is known to remain in the blood for extended periods of time, in proportion to its molecular weight (Pain et al. 1984).

Upon intravenous injection a number of pharmacologically active proteins and peptides have been found to undergo proteolytic degradation or a rapid clearance from the circulation, thereby severely impeding the therapeutic use of these molecules. Various attempts have been made (Marshall et al. 1977, Brucato & Pizzo 1990, Benbough et al. 1979) to prolong the duration of proteins in circulation in order to give them sufficient time to find their target as in site-avoidance targeting. Further, the immune system could also become responsive to repeated administration of an exogenous product. Consequently, the need arises to modify pharmacologically active molecules in order to eliminate such a response, while augmenting the biological property. Conjugation with dextran was found to yield products that could meet the above requirements reasonably well (Melton et al. 1987, Yasuda et al 1990).



The rationale for the approach of coating an enzyme with water-soluble polysaccharides in order to lend it *in vivo* stability, lies in mimicking the lysosomal enzymes, most of which are glycoproteins. It has been suggested that the carbohydrate residues help to maintain the stability of glycoenzymes in the highly acidic environment of the lysosomes (Pazur et al. 1970).

In this review, the effect of chemical conjugation with dextran on the physicochemical and biochemical properties of enzymes and other molecules has been reported. The relevance of the altered properties of the conjugates in context of their prolonged persistence in the circulation and inside the cell has also been discussed. Although a number of hydrophilic polymers such as polyethyleneglycol (PEG) and polyvinylpyrrolidone have been utilized for the purpose of conjugation, we are limiting the scope of this review only to dextran conjugated enzymes.

### PREPARATION OF CONJUGATES

Various methods of conjugation have been extensively reviewed earlier (Mumtaz & Bachhawat 1991). Activation of dextran prior to conjugation generally involves the generation of reactive groups as in periodate oxidation of dextran, or introducing labile groups adjacent to an electrophilic centre, into the dextran backbone. The attacking nucleophile is usually the amino function on the protein to be conjugated. Thus reactive groups can be introduced directly into the dextran molecule by activation with cyanogen halides/organic cyanates, cyanuric chloride, bisoxirane, epichlorohydrin, divinyl sulphone, organic sulphonyl chloride etc. Cyanogen bromide has been the conventional reagent for cyanylation so far, though recently the use of 1-cyano-4-dimethyl-aminopyridinium-tetrafluoroborate has been found to be less aggressive to proteins (Andersson et al. 1991). Other strategies involve the activation of both components of the conjugate. In one such method employed for coupling antibodies to dextran, controlled periodate oxidation of the oligosaccharide residues in the Fc portion of the antibodies was carried out which resulted in the generation of aldehydic functions within the molecule. This activated antibody was then coupled to the hydrazide groups in hydrazido-introduced dextran. Substantial antigen binding properties of the antibody were found to be retained upon conjugation with dextran as revealed by radio immunoassay. This probably a result of coupling to the antibody at a point distal from the antigen binding side (Rakestraw et al. 1990).

On the other hand, conjugates prepared by the Maillard reaction require no prior activation of either component. This involves controlled dry heating of a lyophilized mixture of protein and dextran in which a condensation between the amino groups on the protein and the single reducing terminus of dextran is likely to occur. Thus heat stable enzymes such as lysozyme which retain their catalytic functions even under the above conditions have been conjugated to dextran (Nakamura et al. 1991).

Finally, if the active site of enzyme is altered upon conjugation there may be loss of enzyme activity in the conjugate. Prior incubation of the enzyme with the substrate has been used as an effective method of protecting the active site (Blomhoff & Christensen 1983).

### PHYSICOCHEMICAL PROPERTIES OF CONJUGATES

Changed physicochemical properties such as stability to urea denaturation, increased thermal stability and the capacity to withstand extremes in pH are a reflection of the inherent stability of the conjugates in comparison to their unmodified counterparts. This relative stability of the enzyme in the conjugated form could be a corollary effect of contributions from the following factors.

*Hydrophilicity of the conjugates:* The carbohydrate coating has been suggested to be responsible for preserving a water shield around the protein (Blomhoff & Christensen 1983). Thus when the high capacity water binding salt magnesium sulfate was added at a low concentration to solutions of  $\beta$ -galactosidase and the dextran- $\beta$ -galactosidase conjugate, resistance to thermoinactivation was found to be much higher for the conjugated form. Further, the high thermal stability of the enzyme observed upon conjugation with hydrophilic polysaccharides dextran and acetylated dextran, decreased considerably when hydrophobic methylated dextran was used to modify the enzyme. These observations surely call for a role of hydration in providing protection against thermal denaturation of proteins.

*Intramolecular cross-linking in conjugates:* Intermolecular reaction between protein and dextran during conjugation is believed to result in several linkages at different places (Marshall & Rabinowitz 1976). An analogy has been drawn between the disulfide bridges in proteins and such intramolecular cross linking of enzyme molecules by dextran polymers. Contribution from this factor is considered to be the most important for the conformational stabilization of the tertiary structure of the protein in the

conjugate. Oligomeric enzymes are thus also stabilized against dissociation. Thus, the resistance to thermoinactivation of  $\beta$ -galactosidase conjugate was lost on treatment with dextranase (Blomhoff & Christensen 1983). In another report, removal of calcium ions by a chelating agent resulted in a faster loss of activity for  $\alpha$ -amylase as compared to its conjugated form (Marshall 1978).

*Intermolecular hydrogen-bonds between protein and dextran in conjugates* : Hydrogen-bonds between the hydrophilic polymer dextran and hydrophilic amino acid residues of the protein may lend additional conformational stability to the conjugate (Blomhoff & Christensen 1983). In other words, the extra stability of some equilibrium states of the protein from among several states in dynamic equilibrium, arises from the additional hydrogen-bonding between the two components of the conjugate. The protein in the conjugate is thus left to oscillate between even fewer conformational states. Hence, achievement of additional resistance to unfolding and thereby inactivation by multipoint binding to dextran is the effect of increased rigidity of the protein molecule in the conjugate (Klibanov 1979).

## BIOCHEMICAL PROPERTIES OF CONJUGATES

Biochemical properties such as the affinity of an enzyme for its substrate, the extent of binding of an antigen to its antibody or a ligand to its receptor, may be altered upon conjugation, thereby reflecting the differences in their biochemical properties as compared to the parent proteins (Melton et al. 1987, Yasuda et al. 1990, Andersson et al. 1991 and Rajagopalan et al. 1985). Steric hindrance by the polysaccharide coating to approaching the conjugate is forwarded as one of the primary factors responsible for such observations. Marshall has suggested intermolecular cross linking of polypeptide chains in the resulting macromolecular aggregates to be responsible for sterically preventing the entry of substrate or even inhibitor molecules into the matrix, depending on their molecular size (Marshall 1978). Hence, while caseinolytic activity of trypsin-dextran conjugate was only 7% of its native form, specific activity for autodigestion was 53% of the unmodified enzyme. In the presence of soyabean trypsin inhibitor, native trypsin was left with only 6% of its original activity whereas 29% was found to be retained by the dextran coupled enzyme (Marshall & Rabinowitz 1976).

In this context, it becomes pertinent to address the phenomenon of enhanced resistance to proteolytic cleavage of these conjugates, where the conjugate plays the role of the substrate (Melton et al. 1987, Blomhoff & Christensen 1983). Apart from having to face steric inaccessibility, a proteolytic enzyme has to approach the extremely hydrated surface of the conjugate. Atha and Ingham working with another hydrophilic polymer PEG have commented upon the effect of this structure maker on proteins in solution (Atha & Ingham 1981). According to them, restructuring of water molecules around PEG leads to a steric exclusion of proteins which are thus precipitated by a mechanism that is primarily entropic. In addition, PEG may promote phase separation as in the case of charged hydrophilic proteins, with the driving force of this thermodynamically controlled crystal formation being provided by unfavourable thermodynamic interactions between the solvent system and proteins (Lee & Lee 1981). Such a mechanism may be operative for the hydrophilic dextran coated proteins as well so that any approaching proteins such as proteolytic enzymes, antibodies receptors on cells or circulating opsonins may be prevented by the hydrated barrier from approaching any further.

#### ALTERED SURFACE PROPERTIES OF THE CONJUGATES

*Size of the conjugates :* The urinary recovery of a small molecular weight drug mitomycin C was found to decrease from 23% of the injected dose to almost none when the molecular weight of the covalent attached dextran was increased from 10,000 to 500,000 (Takakura et al. 1987). This observation has been attributed to reduced glomerular filtration as the relative molecular masses of the conjugates is increased.

*Interaction of conjugates with a complementary surface:* Covalent attachment to water soluble supports such as dextran may considerably alter the surface structure of the parent protein in which case there could be a reduced recognition of the conjugate by the corresponding complementary surface.

*Building to receptor:* In an *in vitro* binding assay using cultured glioma cells, maximum binding of conjugated epidermal growth factor (EGF) was found to be at 1 hour as against 20-30 minutes for the native form (Andersson et. al. 1991). Anti-immunoglobulin-dextran conjugate was also shown to bind to the surface immunoglobulin receptor on *B* cells, though the mode of activation of *B* cells by the conjugate was suggested to be different from that of the unmodified anti-immunoglobulin. Not only

could the anti-immunoglobulin coupled to dextran, stimulate a 10 fold greater increase in B cell proliferation as compared to native anti-immunoglobulin, it could also activate the B cells at concentrations about 1,000-10,000 fold lower than those required by the unconjugated form (Brunswick et al. 1990).

In an interesting experiment,  $\beta$ -galactosidase which has been shown to be taken up by the mannose-6-phosphate receptor found in the liver, was conjugated to dextran. The carbohydrate residues reportedly involved in the receptor mediated endocytic uptake of the enzyme were apparently not masked in the conjugate. This is probably due to inadequate shielding of the mannose-6-phosphate residues by the low molecular weight dextran molecules and almost identical uptake rates of both forms by the hepatic cells were observed *in vitro* (Blomhoff et al. 1983).

In our laboratory, horse radish peroxidase (HRP), a mannose containing glycoprotein, was shown to be taken up *in vivo* by a mannose receptor mediated endocytic pathway into the liver, since the uptake was inhibited in the presence of mannan. However, on conjugation with dextran, the liver uptake of HRP was considerably reduced though it was still inhibited to an extent by mannan. In addition, while maximum native HRP was internalized into hepatic cells in less than 30 minutes, the maximum uptake observed for the conjugates was at 1 hour. This reduction of interaction of the mannose residues of HRP-dextran conjugate with the mannose specific receptors has been described to be due to the formation of a steric shield of dextran molecules around the enzymes [unpublished results].

*Binding to antibody:* An increase in molecular weight of dextran conjugated to uricase was found to be accompanied with a simultaneous decrease of reactivity of the conjugate toward antisera (Yasuda et al. 1990). Interestingly, in spite of some loss in receptor recognition observed for HRP-dextran as mentioned earlier, the conjugate exhibited complete retention of its antibody binding properties.

In another report, the major allergen of Ragweed pollen antigen E was coupled to dextrans of two different molecular weights (King et al. 1975). Precipitation reactions of these conjugates with anti-antigen E sera revealed a more pronounced reduction in binding to antibody for the larger sized product while the zone of equivalence was found to shift to a lower concentration of antigen for both the modified antigens. Further, this reduction in binding to antibody could be reversed on dextranase treatment

indicating that it was not a result of denaturation of the antigen while conjugating with dextran.

### ENHANCED CIRCULATORY LIVES OF CONJUGATES

Several enzymes have been shown to exhibit increased lifetimes in circulation upon conjugation with dextran. The circulatory life of  $\alpha$ -amylase for example, was found to increase about five fold when it was covalently modified with dextran. The increase in plasma persistence observed for most dextran conjugates (table 1) may be largely due to the combined effect of enhanced resistance to proteolytic degradation and reduced hepatic uptake of these conjugates.

The decrease susceptibility of enzyme-dextran conjugates to proteolysis has been described by several workers (Marshall & Rabinowitz 1976, Blomhoff et al. 1983). As already mentioned, reasons for such observations *in vitro* have been attributed to the formidable barriers of hydration and steric hindrance afforded by the hydrophilic polysaccharide coating of the conjugates. The same factors are believed to be responsible for similar observations *in vitro*. Further, the enhanced resistance to proteolytic cleavage by enzymes such as trypsin, is suggested to be due to the unavailability of the substrate lysine residues involved in covalent linkage with dextran.

Reduced uptake of the conjugates is probably due to loss in receptor recognition which in turn is a fall out of a altered surface properties of these conjugates described above. Thus while free HRP could barely be detected in the serum after 30 minutes of intravenous injection, about 48% of the administered dose of the HRP-dextran conjugate could still be recovered at one hour. In contrast only a slight difference in plasma clearance rates of native and conjugated  $\beta$ -galactosidase were observed with both exhibiting half lives of about 1 minute (Blomhoff et al. 1983). These observations are therefore in accordance with the *in vitro* results mentioned above which demonstrate identical uptake rates of  $\beta$ -galactosidase and  $\beta$ -galactosidase-dextran conjugate by cultured hepatic cells.

The hydrophilic coating of the conjugated enzyme may prevent its association with circulating opsonins. Adsorption of opsonins which seek hydrophobic surfaces is believed to occur upon intravenous injection of a foreign body (Moghimi & Patel 1988, 1989). In the presence of preexisting antibodies specific for the injected enzyme, binding to the

antibody may lead to the clearance of the foreign protein from circulation. In view of the altered properties of dextran-enzyme conjugates however, such a neutralization by circulating antibodies and association with opsonins would be difficult.

The reduced hepatic uptake of conjugated enzymes is also believed to be a result of decreased cationic charge of covalently modified enzymes which accompanies the modification of the surface lysine residues. As a result the circulatory life of the conjugates is prolonged. Further, glomerular filtration of small proteins and peptides is believed to be the underlying reason for their short circulatory lives (Venkatachalam & Rennke 1978). The increase relative mass of the conjugated enzyme has been suggested to be responsible for a decrease in urinary excretion and thereby enhanced circulatory life.

The increased circulating times may also be due to the reduced immunogenicity exhibited by the dextran coated enzymes. The antibody titre of uricase conjugated to dextran for instance, was found to be decreased, especially for the higher molecular weight dextran. Further, the antibodies raised against either form i.e. native or conjugated protein, could interact with the conjugated uricase *in vitro*, to a much lesser degree (Yasuda et al. 1990). Thus *in vivo*, reduced recognition by the immune system could lead to a decrease in immunogenicity of the covalently modified enzyme.

**Table 1**  
*Circulatory half lives of dextran conjugates*

| Enzyme                 | Half Life<br>of Enzyme | Half Life of<br>Conjugate | Reference                  |
|------------------------|------------------------|---------------------------|----------------------------|
| Carboxypeptidase-G2    | 3.1 hr                 | 45.6 hr                   | Melton et al.<br>(1987)    |
| Uricase                | 0.6 hr                 | 7.5 hr                    | Yasuda et al.<br>(1990)    |
| Haemoglobin            | 1.5 hr                 | 2.5 hr                    | Tam et al.<br>(1976)       |
| Asparaginase           | 11.0 hr                | 190.0 hr                  | Benbough et al.<br>(1979)  |
| Catalase               | 17 min                 | 140 min.                  | Marshall et al.<br>(1977)  |
| $\beta$ -galactosidase | 1 min.                 | 1 min.                    | Blomhoff et. al.<br>(1983) |

Hence, reduction in glomerular filtration, decreased immunogenicity, reduced recognition by receptors on cells and circulating antibodies and reduced interactions with opsonins and proteolytic enzymes are the various reasons that have been advanced to explain the enhanced circulatory lives of dextran coated macromolecules (table 1).

### INTRACELLULAR STABILITY OF CONJUGATES

Internalization of ligand via receptor mediated endocytosis could lead to a rapid degradation of the ligand. In an *in vitro* study, an increased stability against intracellular degradation has been demonstrated for the radiolabelled dextran-  $\beta$ -galactosidase conjugate (Blomhoff et al. 1983). As compared to the free enzyme, degradation was found to be reduced by 35% in hepatocytes and by 43% in non-parenchymal cells after 80 minutes and 40 minutes of incubation, respectively. Such observations surely are a fall out of the acquired intrinsic structural stability of the chemically engineered enzyme. In this study however, no attempts were made to check the extent of retention of the immunological and catalytic activities of the covalently modified enzyme inside the cell. Hence it is not possible to evaluate the intracellular intactness of the modified  $\beta$ -galactosidase.

In another *in vitro* study (Andersson et al. 1991) the fate of the ligand EGF was found to change drastically upon covalent modification with dextran, from the view point of stability against degradation. Radioactivity of the labelled EGF was found to remain in the cell for more than 20 hours. In contrast a rapid disappearance of native EGF was observed within a few hours. In this case also, the retention of the immunological and biological property of the modified ligand was not determined upon internalization.

While investigating the *in vivo* fate of the radiolabelled carboxypeptidase-G2 conjugate, significant retention of the conjugated enzyme could be detected (Melton et al. 1987) with about 11.7% and 3.2% of the injected dose being recovered from the liver after 12 hours and 48 hours, respectively. In addition when the radiolabel was located on the dextran part of the conjugate it was found to disappear at a much slower rate than when it was present on the enzyme component. Hence it was concluded that of the two components, dextran was being degraded and released at a much slower rate. A comparison of the stability of the conjugated and native forms was not possible since native carboxypeptidase-G2 was found to demonstrate insignificant tissue



uptake. Further, the radiolabelling experiments conducted by these workers as well do not lend any insight into the extent of stabilization of the immunological property and catalytic function of the modified enzyme inside the cell.

In our laboratory increased intracellular stability of an enzyme (HRP) conjugated to dextran has been demonstrated for the first time *in vivo*. For a direct comparison of the intracellular persistence time of native and modified HRP, it was necessary to localize equal amounts of both forms of the enzyme intracellularly. However, as already mentioned the rate of uptake of HRP was found to be considerably inhibited upon conjugation with dextran. Therefore, in order to deliver the native and the modified enzyme at the same rate to the liver, both forms were encapsulated into liposomes and injected intravenously. This resulted in the accumulation of equal amounts of both forms in the liver at 30 minutes. As much as 40% of the dextran-HRP conjugate accumulated in the liver within 30 minutes continued to be localized even after 24 hours. In contrast to this no HRP could be detected in the liver after the same period of time. An analogous profile was observed in the mitochondrial/lysosomal fraction, obtained by a subcellular fractionation of the total liver homogenate. Almost 70% of the conjugate accumulated at 30 minutes in the lysosomal fraction could be recovered after 24 hours as compared to 5% of the unconjugated HRP. The intactness of the enzyme *in vivo* was demonstrated by measuring its catalytic activity and capacity to bind to antibody. This is in contrast to the radiolabelling methods used by the earlier workers which do not throw any light upon the stabilization of the enzymes inside the cell from the angle of preservation of enzyme activity (Mumtaz & Bachhawat 1992, in Press).

### THERAPEUTIC POTENTIAL OF DEXTRAN CONJUGATES

Immunochemical studies with the Regweed pollen allergen, antigen-E conjugated with dextran have indicated reduced allergenicity of the conjugate (King et al. 1975). The reduced allergic response is believed to be due to the availability of a lesser number of antigenic determinants in the coupled form. The extent of binding of the conjugate to the surface bound receptor of the IgE class which require multivalent antigens is therefore reduced. This property of reduced allergenicity observed for some dextran conjugated enzymes have tremendous therapeutic potential.

Prolonged circulating times are essential for certain enzymes where the efficacy of the enzyme depends upon its retention in circulation. Covalent modification of the anti-leukaemia drug L-asparaginase with dextran was shown to extend its circulatory half life (Benbough et al. 1979), thereby enhancing its therapeutic potential.

One of the strategies for the treatment of lysosomal storage disorder has been enzyme replacement therapy (Colley & Ryman, 1976). Owing to the short intracellular half lives of therapeutically active enzymes such strategies have had limited application. This approach of covalently modifying enzymes with hydrophilic polymers such as dextran, has been extended for the stabilization of enzymes in the lysosomes (Blomhoff et al. 1983). In our laboratory, preservation catalytic activity of HRP-dextran conjugate, has been amply demonstrated, thereby proving conclusively that dextran coupled enzymes are much better equipped to deal with the harsh environment of the lysosome within the cell. The *in vivo* stability of the proteins before and after modification with dextran is depicted here in the form of a cartoon.

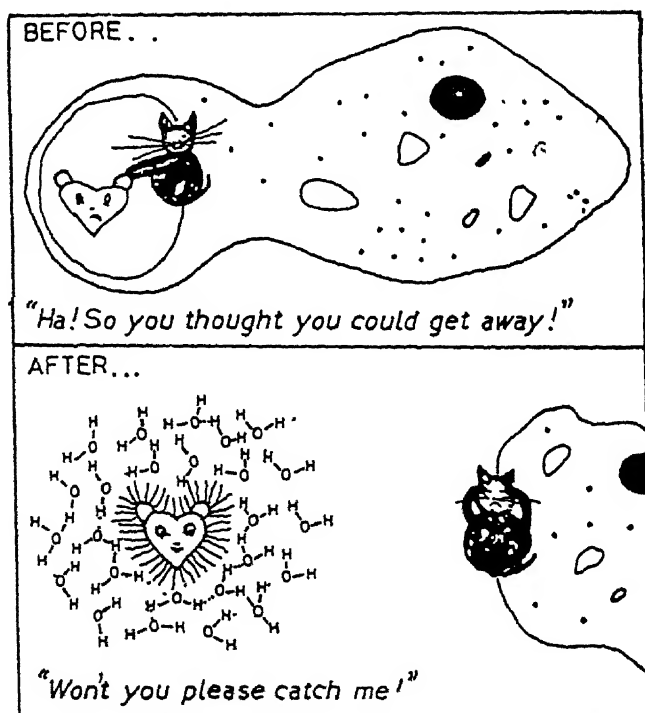


FIG.

## ACKNOWLEDGEMENTS

G. M. is a Senior Research Fellow of the Council of Scientific and Industrial Research, S.M. is a Senior Research Fellow of University Grants Commission. B.K.B. is a Bhatnagar Fellow of the Council of Scientific and Industrial Research. We are grateful to Ms. Poonam for preparing the manuscript on computer. Financial assistance from the Department of Biotechnology is gratefully acknowledged.

## References

1. Andersson A, Holmberg A, Carlsson J, Ponten J and Westermarck B 1991 Binding of epidermal growth factor-dextran conjugates to cultured glioma cells; *Int J. Cancer* **47** 439-444
2. Atha D H and Ingham KC 1981 Mechanism of precipitation of proteins by PEG, *J Biol. Chem.* **256** 12108-12117
3. Benbough JE, Wiblin CN, Rafter TNA and Lee J 1979 The effect of chemical modification of L-asparaginase on its persistence in circulating blood of animals; *Biochem. Pharmacol.* **28** 833-839
4. Blomhoff HK and Christensen TB 1983 Effect of dextran and dextran modifications on the thermal and proteolytic stability of conjugated bovine testis *b*-galactosidase and human serum albumin; *Biochem. Biophys. Acta* **743** 401-407  
– Blomhoff R and Christensen TB 1983 Enhanced stability of  $\beta$ -galactosidase in parenchymal and non-parenchymal liver cells by conjugation with dextran; *Biochem. Biophys. Acta* **757** 202-208
5. Brucato FH and Pizzo SV 1990 Catabolism of streptokinase and polyethylene glycol-streptokinase: Evidence for transport of intact forms through the biliary system in the mouse; *Blood* **76** 73-79
6. Brunswick M, Bonvini E, Francis M, Felder CC, Hoffman T and Mond J 1990 Absence of demonstrable phospholipid turnover in B cells stimulated by low mitogenic concentrations of dextran-anti-immunoglobulin conjugates; *Eur J Immunol.* **20** 855-861
7. Colley CM and Ryman BE 1976 The use of a lyposomally entrapped enzyme in the treatment of an artificial storage condition; *Biochem. Biophys. Acta* **451** 417-425
8. King TP, Kochoumian L, Ishizaka K, Lichtenstein LM and Normal PS 1975 Immunochemical studies of dextran coupled ragweed pollen allergen, antigen-E; *Arch. Biochem. Biophys.* **169** 464-473
9. Kilibanov AM 1979 Enzyme stabilization by immobilization; *Anal. Biochem.* **93** 1-25

10. Lee JC and Lee LLY 1981 Preferential solvent interactions between proteins and PEG; *J. Biol. Chem.* **256** 625-631
11. Marshall JJ 1978 Manipulation of the properties of enzymes by covalent attachment of carbohydrate; *Trends. Biol. Sci.* **3** 79-83
 

— — Humphreys JD and Abramson SL 1977 Attachment of carbohydrate to enzyme increases their circulatory lifetimes; *FEBS Lett.* **83** 249-252

— — and Rabinowitz ML 1976 Preparation and characterization of a dextran-trypsin conjugate; *J. Biol. Chem.* **251** 1081-1087
12. Melton RG, Wiblin CN, Foster RL and Sherwood RF 1987 Covalent linkage of carboxypeptidase-G<sub>2</sub> to soluble dextrans-1: Properties of conjugates and effects on plasma persistence in mice; *Biochem. Pharmacol.* **36** 105-112
 

— —, — —, Basterville A, Foster RL and Sherwood RF 1987 Covalent linkage of carboxypeptidase-G<sub>2</sub> to soluble dextrans-11; *In vivo* distribution and fate of conjugates; *Biochem Pharmacol.* **36** 113-121

— — and Patel H M 1988 Tissue specific opsonins for phagocytic cells and their different affinity for cholesterol-rich liposomes; *FEBS Lett.* **233** 143-147
13. Moghim SM and Patel 1989 Serum opsonins and phagocytosis of saturated and unsaturated phospholipid liposomes; *Biochem. Biophys. Acta* **984** 384-387
14. Mumtaz S and Bachhawat BK 1991 Conjugation of proteins and enzymes with hydrophilic polymers and their applications; *Indian J. Biochem. Biophys.* **28** 346-351
 

— — — — 1992 Enhanced intracellular stability of dextran horse radish peroxidase conjugate: an approach to enzyme replacement therapy; *Biochem. Biophys. Acta* (In Press)
15. Nakamura S, Kato A and Kobayashi K 1991 New antimicrobial characteristics of lysozyme-dextran conjugate; *J Agric. Food Chem.* **39** 647-650
16. Pain D, Das PK, Ghosh P and Bachhawat BK 1984 Increased circulatory half-life of liposomes after conjunction with dextran; *J. Biosci.* **6** 811-816
17. Pazur J H, Knull H R and Simpson D L 1970 Glycoenzymes: A note on the role for the carbohydrate moieties; *Biochem. Biophys. Res. Comm* **40** 110-116
18. Rajagopalan S, Gonias SL and Pizzo SV 1985 A nonantigenic covalent streptokinase-polyethylene glycol complex with plasminogen activator function; *J Clin. Invest.* **75** 431-419
19. Rakestraw SL, Tompkins RG and Yarmush ML 1990 Antibody-targeted photolysis: *In vitro* studies with Sn(IV) chlorine e6 covalently bound to monoclonal antibodies using a modified dextran carrier; *Proc. Natl. Acad. Sci. USA* **87** 4217-4221

20. Takakura Y, Atsumi R, Hashida M and Sezaki H 1987 Development of a novel polymeric prodrug of mitomycin C, mitomycin c-dextran conjugate with anionic charge.II. Disposition and pharmacokinetics following intravenous and intramuscular administration; *Int. J. Pharmaceut.* **37** 145-154
21. Tam SC, Blumenstein J and Wong JTF 1976 Soluble dextran-hemoglobin complex as a potential blood substitute; *Proc. Natl Acad. Sci USA* **73** 2128-2131
22. Venktachalam MA and Rennke HG 1978 The structural and molecular basis of glomerular filtration; *Circulation Res* **43** 337-347
23. Yasuda Y, Fujita T, Takakura Y, Hashida M and Sezaki H 1990 Biochemical and biopharmaceutical properties of macromolecular conjugates of uricase with dextran and polyethylene glycol; *Chem. Pharm Bull* **38** 2053-2056



**Raja Ramanna** (b. 28 January 1925) did Ph.D. (1948) from University of London, UK; He has received D.Sc. (h.c.) from several universities and Honorary Degree of Desikottama (D.Litt.), from Visva Bharati. He is Director, National Institute of Advanced Studies, Indian Institute of Science (IISc) Campus, Bangalore (1987- Jan 1990, Dec 1990- ), Chairman, Governing Council, IISc, Bangalore.

Ramanna's significant research achievements are in the field of design and commissioning of research reactors, experimental and theoretical investigations of neutron thermalization and neutron scattering, investigations on fission, including a stochastic theory of the fission process, a geometric theory of atomic and nuclear structure, nuclear exploration for civilian applications, planning accelerator programmes, fast-reactor development and defence-oriented research programmes. More specifically, Ramanna collaborated in the design, installation and commissioning of the research reactors Apsara and Cirus and the variable energy cyclotron (Calcutta) and fast-breeder test reactor (Kalpakkam). Led the group which was responsible for the nuclear explosion (Pokharan, 1974) aimed at exploiting the atom for socio-economic uses. He has written autobiography 'Years of Pilgrimage' (1991), 'The Structure of Music in Raga and Western Systems' (1993).

Ramanna is Fellow, Indian Academy of Sciences (Vice-President, 1977-78) and Indian Society of Engineers; President, Indian Physics Association; Honorary Fellow, IISc and TIFR; President, INSA (1977-78). He is the recipient of Shanti Swarup Bhatnagar Prize (1963); Meghnad Saha Medal (INSA) (1984); Pandit Jawaharlal Nehru Medal (Madhya Pradesh Government) (1983); Om Prakash Bhasin Award (1985); R.D. Birla Memorial Award (Indian Physics Association) (1985-86); Rajyotsava State Award (Karnataka Government); President, 30th General Conference, IAEA; Padma Shri (1968); Padma Bhushan (1973); Padma Vibhushan (1975).

*Raja Ramanna was elected to the fellowship of the Academy in 1973.*

**Address:** Director, National Institute of Advanced Studies, Indian Institute of Science Campus, Bangalore 560 012; Tel: (080) 344351, 346594, 346251; Fax: 91-80-346634; Res: 407, II Block, 9th Cross Road, Rabindranath Tagore Nagar, Bangalore-560 032; Tel: (080) 330815, 333284.

## A 15-YEAR PROGRAMME FOR NUCLEAR POWER IN INDIA

RAJA RAMANNA

### PREAMBLE

*I feel deeply honoured that the Indian National Science Academy thought it fit to award the Meghnad Saha Medal to me for the year 1984. Though I had met Prof. Saha on various occasions, I had known him only to the extent that he could just recognise me. We are all aware of his immense contributions to Physics, but few know of the deep interest taken by him in the area of national planning. He was the Chairman of the National Planning Committee of the Indian National Congress during the 1940's and was always interested in the long term economic development of the country. I have, therefore, chosen to speak about planning in an area which affects the entire spectrum of the economic development of our country—viz. planning for Electric Power Generation. I will concentrate on the development of nuclear power and in particular on our experience and our confidence in embarking on a 15-year programme for nuclear power which will, to a great extent, meet the steeply rising demand for electricity in the years to come.*

### INTRODUCTION

Ever since the setting up of the Atomic Energy Establishment in the mid-1950s, one of the primary objectives has been to achieve self sufficiency in the technology of harnessing the power of the atom to meet the country's growing energy requirements. The demand for electrical energy has been growing exponentially over the years and projections indicate that India will have to nearly treble its present electrical generating capacity by the turn of century. The magnitude of this, demands that all proven and available technologies be fully exploited. Paucity of natural non-renewable resources like oil or high-grade coal and restrictions on the availability of hydro-potential, lead to the conclusion that these resources alone cannot meet the projected demand for electric power. Other exotic technologies being talked of, are not feasible—either technically or economically. The self-sufficiency attained in the field of nuclear power with the commissioning of the Madras Atomic Power Station (first unit)

last year, now enables this source of electric power to contribute significantly in meeting the energy requirements during the years to come.

### ENERGY—THE SOURCE OF ECONOMIC GROWTH

It is an accepted fact that the per-capita consumption of *total energy* of any country has a direct correlation to its Gross Domestic Product and is thus a measure of its present stage of economic development. Presently, nearly 40 per cent of India's total energy requirements is met from non-commercial sources like firewood, agricultural waste, cowdung, animal power and manual labour. In fact, the demand for firewood has been estimated to denude the existing forest wealth at the rate of 10,000 square kilometres every year resulting in very severe ecological imbalance in the country: according to some ecological experts only 10 per cent of the land is now estimated to be covered with forests compared to 60 per cent about 100 years back.

The remaining 60 per cent of the total energy requirements is classified as commercial energy, of which the contribution of electricity is estimated at around 30 per cent i.e. just about 18 per cent of the total energy needs. Hence, the validity of a 'direct' correlation between per-capita electricity consumption and the prosperity of the country may not be universally valid. On the other hand, the industrial and agricultural sectors contribute most to the economy and consume over 75 per cent of the total electricity generated. In this way, per capita electricity consumption does reflect to a large extent, the economic growth of the country.

### RATE OF GROWTH OF ELECTRIC GENERATION AND PROJECTED DEMAND

By any standard, the growth of the power sector in India has been quite impressive as can be seen from Fig. 1.

Even though the electrical energy generated every year has multiplied nearly thirty-fold during the last three decades, the country's per capita commercial energy consumption is less than a tenth of the world average. Not surprisingly our economic growth has resulted in the demand for electricity outstripping its availability. Today the shortfall is estimated to be nearly 10 to 15 per cent. To some extent, it is due to the low plant utilisation factors and high transmission and distribution losses.



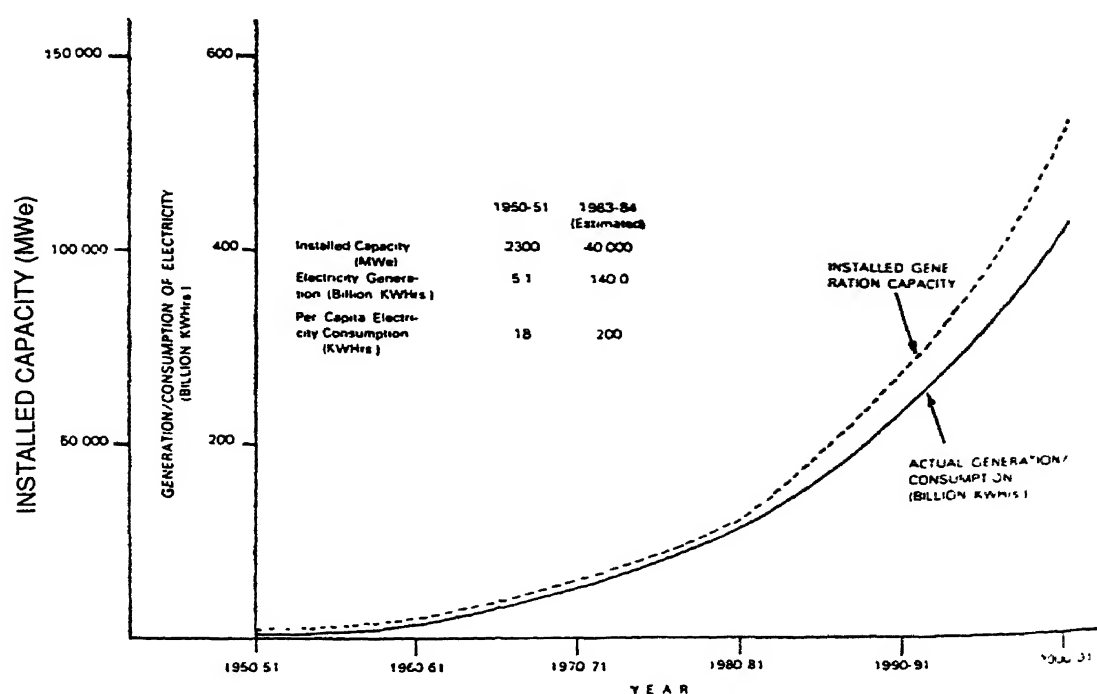


FIG 1 Past and projected growth of electric power generation capacity and actual generation/consumption of electric power in India

Various projections that have been made indicate that the installed generating capacity required by the year 2000 A.D. should be 120,000 to 150,000 MWe. If one also takes into account replacement of ageing power plants, the *gross additions* to the present generation capacity should be 100,000 to 120,000 MWe. The need for nearly trebling the installed electrical generating capacity within the next two decades, necessitates an examination of the various sources of energy and of the alternatives available for tapping them on a commercial scale for generating electricity.

### SOURCES OF ELECTRICAL ENERGY

In recent years, various suggestions concerning alternative sources like solar, geothermal, tidal, etc. have been made as feasible solutions towards meeting the projected demand for electrical energy. These require a critical examination.

Primarily, any energy source can be classified as: Renewable or Non-Renewable. Renewable sources comprise of — Solar, Wind, Geothermal, Tidal, Ocean Thermal and Hydro, and Non-Renewable sources comprise of Coal, Oil, Gas and Nuclear. Logic dictates the use of

renewable sources to the maximum extent and the conservation and most efficient use of the scarce non-renewable sources.

A comparison of the energy potential of some of these sources is given in Table I. The energy content and the temperature at which a source is available will determine the theoretical limits of tapping them for electric power. The technology available and its economics will then determine its feasibility for being taken up on a commercial scale.

**Table I**  
*Comparison of potential of various energy sources*

| Energy source   | Magnitude of energy content (Kw Hrs.) it is available | Max. temperature levels at which (°Kelvin) |
|---|---|--|
| Fission energy per cubic metre of Uranium-235                 | $4 \times 10^{11}$                                    | $1 \times 10^{11}$                         |
| Fusion energy per cubic metre of deuterium                    | $1 \times 10^7$                                       | $1 \times 10^{10}$                         |
| Fossil fuel energy per cubic metre of fuel (oil or coal)      | $1 \times 10^4$                                       | $1 \times 10^4$                            |
| Hydro energy per cubic metre of water (30 metres head)        | 1   | Ambient                                    |
| Solar energy per hour per square metre                        | 1   | $6 \times 10^3$                            |
| Wind energy (from a 36 Kmphr wind) per hour per square metre  | $5 \times 10^{-1}$                                    | Ambient                                    |
| Geothermal energy (earth's average) per hour per square metre | $6 \times 10^{-5}$                                    | $5 \times 10^2$                            |

## RENEWABLE SOURCES OF ENERGY

### *Solar Energy*

Though solar energy is generally perceived as just sun-light, strictly speaking it governs the availability of other renewable sources like wind, tidal, hydro and ocean thermal. By heating the earth and its atmosphere, the sun is actually responsible for generating wind, waves, rainfall and ocean temperature gradients. Primary solar energy can be converted to electricity, in two ways: one, by direct conversion using photovoltaic solar cells and the other indirectly, by concentrating the solar energy to produce steam and then running an electric generator. A typical 8 cm. solar cell (assuming a maximum efficiency of 16 per cent) can yield about half a watt and is also expensive. The solar-thermal route requires large open

areas for locating precisely controlled reflectors/collectors to 'follow' the sun.

The overall annual performance in both these methods could vary by as much as 20 per cent due to changing weather and atmospheric conditions. The fact that the sun is available for only a portion of the day necessitates back-up batteries and storage devices. In addition, the low energy content of this source combined with the low efficiency of the conversion systems dictate that solar energy for generating electricity on a commercial scale will not be feasible under the present circumstances. It will be valuable for rural electrification in remote locations, where transmission of power would be costly. It would also be of use in specific applications where small quantities of electricity are needed.

### *Wind Energy*

The technology of wind mills coupled to electric generators is a proven one. Average wind speeds in India are generally in the 8 to 10 kmph range and even these are generally found in a few coastal and hilly areas. The energy potential of this source is about half that of solar and consequently the size of the wind mills would have to be gigantic to generate even a few hundred kilowatts of electric power. For example, even with a wind speed of 70kmph, a 300kw wind — electric unit would require a 22-metre diameter rotor with three blades, and a 22-metre tall tower. Thus, wind energy will not be able to contribute significantly to electricity generation in the near future.

### *Geothermal, Tidal and Ocean Thermal*

Theoretically all these are possible sources of energy, but are constrained by their locations and limited potential. Utilisation of geothermal energy is being investigated at locations near hot springs. Further development of technology is necessary to tap this source of energy from rocks deep below the earth.

Tidal energy potential is restricted to one or two locations and is perhaps uneconomic to exploit. As the oceans absorb nearly 75 per cent of the incident solar energy, ocean thermal energy conversion (OTEC) appears promising on a long-time fame but the technology for tapping the thermal energy from the oceans has yet to be developed to overcome the associated corrosion and structural problems. The projected capital costs for OTEC are still too high. Though small scale pilot projects are being planned, the first prototypes are expected only after the year 2000. It is

unlikely that this source will be able to contribute to commercial generation of electricity till atleast the middle of the next century.

### *Hydro-power*

The only available renewable source of energy for bulk electricity generation is hydro-power and this source must be exploited to the maximum possible extent. It is cheap, non polluting, and the technology for its exploitation is available in the country. Unfortunately, hydro-plants can be constructed only at specific locations and involve high capital costs and long gestation periods resulting from their being multi-purpose hydel-cum-agricultural schemes. Of the estimated 75,000 MWe of hydropotential in India, about 13,000 MWe (only 17 per cent) has been exploited so far. Plans are under consideration for increasing this to 21,000 to 40,000 MWe (30 per cent to 55 per cent) by the 1990's.

## NON-RENEWABLE SOURCES OF ENERGY

Amongst the non-renewable energy sources, the most scarce resources are oil and gas. Despite an impressive increase in indigenous production of oil, the country will have to continue to import it for sometime to come. India's annual requirement of oil is expected to cross 90 million tonnes by the turn of the century. As against this, if the present oil production rate is stepped up to meet the target of 35 million tonnes per year by 1990 and 60 million tonnes per year by the year 2000, the 'currently recoverable' reserves will be exhausted by then.

In order that this scarce resource is utilised most efficiently, the transportation sector (which already consumes about 47 per cent) would continue to have high priority and would be followed by the petro-chemical and domestic sector which requires kerosene for cooking and lighting. On these grounds, the use of oil for generating electricity has to be ruled out, though a marginal quantity will continue to be required for use as 'flame-stabilisers' in coal-based thermal power plants due to the poor quality of coal available to them.

## OPTIONS AVAILABLE FOR GENERATION OF ELECTRICITY

The only non-renewable sources available for bulk generation of electricity on a 15 to 20 year time-frame will, therefore, be coal and nuclear energy.

## COAL-BASED POWER STATIONS

Though the use of coal for generating electricity is expected to continue increasing steadily, there are several constraining factors like: quantum of proven reserves, quality of coal, logistics of transportation and environmental pollution.

Despite the much quoted statement that our coal reserves are adequate and abundant, a closer examination reveals that of the 112 billion tonnes of coal reserves, only about 26 billion tons come under the 'proven category'. The rest is equally split between "inferred" and "indicated categories". Of the entire reserves, about 24 billion tonnes is good coking grade coal, which is required for use in metallurgical industries. The quality of coal which is available for power generation is extremely poor, sometimes containing upto 40 per cent non-combustibles. The Working Group on Energy Policy has projected a target of 50,000 MWe of thermal power, i.e. over 50 per cent of the total generation capacity, by the turn of the century. This will require nearly 160 million tonnes of coal every year, which would constitute about 37 per cent of the then projected annual production of 430 million tonnes. It will become necessary for much of this thermal power to be located near the pit heads, to minimise the logistics of transportation. A 1000 MWe thermal station operating at 60 per cent capacity factor needs over 10,000 tonnes of coal daily which corresponds to a delivery of 5 to 7 train loads every day.

The environmental aspects of using coal also require special consideration. For example, a 1000 MWe thermal station burning coal with 40 per cent ash content will produce nearly 35 million tonnes of ash during its operating life of 25 years. Many people would be surprised to know that residual coal ash is also a source of radiation. Studies by the Bhabha Atomic Research Centre have shown that the gamma activity in residual coal ash (due to the presence of elements like Radium-226) is higher than that from an operating nuclear power station. Besides, a 100 MWe thermal plant will discharge 40 to 80 tonnes of sulphur dioxide a day. Apart from leading to acid-rain, this is a serious health hazard and expensive equipment will be required to keep such environmental pollution under control.

## THE NUCLEAR OPTION

It is against this background, that a nuclear power programme of 10,000 MWe by the year 2000 has been prepared. It is based on experience gained

to date in the entire spectrum of the nuclear power cycle, from exploration and mining of uranium, to design, construction and commissioning of nuclear power stations — *entirely on our own*. The commissioning of the first unit of the Madras Atomic Power Station is the culmination of all the indigenisation efforts and provides the confidence to take up such a large nuclear power programme.

### NEED FOR AN OPTIMAL MIX OF POWER STATIONS

It needs to be emphasised that nuclear power is not meant to substitute either coal or hydel stations. The earlier concept of 800 km break-even point for nuclear versus thermal power costs was postulated when coal costs were extremely low. It is not entirely valid today. Nuclear power costs will become comparable with thermal power costs even at the pit heads, if past trends of increasing coal prices continue to persist. The relatively low levels in the wages of the labour in coal mines are bound to increase thereby resulting in higher coal prices. However, these costs are just one of the factors to be considered in the overall long term strategy of planning for electric power. A needless argument of nuclear versus thermal power has often been raised in the past. A long term integrated approach to energy planning necessitates an optimal mix of technically and commercially viable options for generating electricity.

### NEED FOR A LONG TERM APPROACH IN PLANNING FOR POWER

If electric power demand in the next 15 to 20 years has to be necessarily met by an optimal mix of hydel, thermal and nuclear power, the inherent characteristics of each of these point to the need for evolving a long term power programme. Hydel plants being multi-purpose schemes have long construction periods extending upto 10 years. Though thermal plants may be constructed within 4/6 years, it would take 8/9 years to explore and develop a coal deposit as it is in the case of exploitation of uranium deposits. Nuclear power plants are expected to have a construction time of 7/8 years. These times clearly indicate that decisions taken today will determine the availability of electric power by the end of the century. It is, therefore, imperative that firm perspective plans for the power sector are drawn up and, it is in this context, the Department of Atomic Energy has proposed a 15-year Nuclear Power programme, which envisages an installed capacity of 10,000 MWe by the year 2000.

Before going into the details of the programme, it would be relevant to examine two questions which are often asked:

- What are the nuclear power programmes of other countries?
- What has been our experience in nuclear power?

## NUCLEAR POWER PROGRAMMES IN OTHER COUNTRIES

A question that is frequently raised is "Since even developed countries are curtailing their nuclear power programmes, is it wise for India to invest heavily in this area?". Such a question is best answered by examining the facts as they exist today.

At the moment, there are over 313 nuclear reactors in operation all over the world contributing about 12 per cent of the total electrical power being generated. Together, they constitute an installed capacity of about 1,90,000 MWe — which is nearly five times India's *present total* installed capacity. In addition, over 200 nuclear power reactors presently under construction are expected to add a further 1,90,000 MWe.

The nuclear power programme in some countries can be seen in Table II. In the U.S.A. for example, where the alleged curtailment of nuclear power plans has received much publicity, 13 per cent of its present generation capacity is from 80 operating nuclear power units having a total capacity of over 63,000 MWe. They have under construction another 50 units with a capacity of over 55,000 MWe. By the year 1990, nuclear power alone will account for nearly 1,20,000 MWe in the U.S.A. In fact, of late, much concern has been expressed in that country about the poor rate of growth of nuclear power in comparison to programmes elsewhere: France is an outstanding example, where 37 per cent (27,000 MWe) of its present generation capacity is from 36 nuclear reactors. In June last year, it became the first country in the world to generate more than half its electricity from nuclear energy. Another 25 reactors are under construction, which will almost double its nuclear power capacity to 56,000 MWe. By 1995, nuclear power in France will account for 75 per cent of the total electrical generation capacity.

**Table II**  
*Nuclear power programmes of various countries*

| Country                            | Status as of end 1983    |  |  |                                |  | Future programmes                        |  |
|------------------------------------|--------------------------|--|--|--------------------------------|--|--|--|
|                                    | No of units in operation | Total Capacity of units in operation (MWe) | Estimated share of nuclear power capacity (as a %age of total electrical capacity) | No of units under construction | Total city of units under construction (MWe) | Projected nuclear capacity (by year) MWe | Projected share of nuclear power capacity (as a %age of total elec capacity) |
| Belgium                            | 6                        | 3,473                                      | 27%  | 2                              | 2,012  | 5,466(1990)                              | 40%  |
| Bulgaria                           | 4                        | 1,632                                      | *  | 2                              | 1,906  | 4,760 (1990)                             | 35%  |
| Canada                             | 15                       | 8,303                                      | 8%   | 8                              | 5,925  | 14,615(1995)                             | 12%  |
| France                             | 36                       | 26,903                                     | 37%  | 25                             | 29,200                                       | ** (1995)                                | 75%  |
| Germany, Democratic Republic       | 5                        | 1,694                                      | *  | **                             | **   | 9,000(1995)                              | **   |
| Germany, Federal Republic          | 16                       | 11,110                                     | 13%  | 11                             | 11,908                                       | 30,000(2000)                             | **   |
| Hungary                            | 1                        | 408  | 7%   | 3                              | 1,224  | 4,760(2000)                              | 45%  |
| India                              | 5                        | 1,095                                      | 3%   | 5                              | 1,175  | 10,000(2000)                             | 10%  |
| Japan                              | 28                       | 19,023                                     | 12%  | 10                             | 10,022                                       | 62,000(2000)                             | 27%  |
| Republic of Korea                  | 3                        | 1,789                                      | 15%  | 6                              | 5,474  | 7,616(1990)                              | 41%  |
| Spain                              | 6                        | 3,760                                      | *  | 9                              | 8,369  | 7,700(1990)                              | **   |
| Sweden                             | 10                       | 7,355                                      | 24%  | 2                              | 2,100  | 9,455(1990)                              | 28%  |
| Taiwan                             | 4                        | 3,110                                      | 27%  | 2                              | 1,814  | 9,928(2000)                              | 41%  |
| Union of Soviet Socialist Republic | 43                       | 20,671                                     | 7%   | 41                             | 38,001                                       | 58,672 (1990)                            | **   |
| United States of America           | 80                       | 63,315                                     | 13%  | 50                             | 55,738                                       | 119,053(1990)                            | **   |
| United Kingdom                     | 35                       | 8,304                                      | 10%  | 7                              | 4,252  | **                                       | **   |

In Bulgaria, Germany (Democratic Republic) and Spain, the electricity generated by nuclear power stations as a percentage of total electricity generated during the year was about 32%, 12% and 9% respectively

\*\* Data not available.

Source - IAEA annual report for 1983 and news releases by the atomic industrial forum.

## OUR EXPERIENCE IN NUCLEAR POWER

Much has been written about slippages, cost escalations and operational problems of nuclear projects in this country. It is unfortunate that this has not been done keeping the overall perspective in mind. Firstly, nuclear power know-how involves high technology and require heavy financial investments in the initial stages. At present it is a monopoly of a few developed countries, which have invested massive amounts over the years



to build up their nuclear industries. Given the primary objective of self-reliance and coupled with the limited financial resources, it was inevitable that the initial steps would be relatively slow. It cannot be denied that there have been some mistakes, but on the whole the achievements are by no means insignificant and certainly a matter of great pride. With the commissioning of the Madras Atomic Power Station last year, India became the sixth country in the world which can design, construct, commission and operate a nuclear power station *all on its own*. During the course of implementation of the 10,000 MWe Nuclear Power programme, India could be in the position to consider sharing this know-how with other developing countries. All the inputs for the operation of nuclear power plants – like heavy water, fuel, high quality fabrication and testing, fuel reprocessing and waste management – are all available today within the country. The progress and status in each of these sectors are as follows.

## NUCLEAR POWER STATIONS

### *Strategy Adopted*

The far sighted decision over two decades back, to opt for the natural uranium fuelled Pressurised Heavy Water type of Reactors (PHWR) instead of the then popular Boiling Water/Pressurised Water Reactors (BWR/PWR) using enriched uranium, has probably contributed most to the self-reliance which India has acquired in the field of nuclear power generation. The decision was an optimal one, in that, it took into account the objective of self-reliance and the country's financial limitations. The PHWR design had the advantages of:

- Possibility of a larger number of Indian sources for the supply of equipment/materials,
- Obviating the need for imported enriched uranium fuel,
- Built-in Safety features,
- Smaller requirements of uranium per MWe installed, and
- Higher plutonium production per MWe installed.

### *Tarapur Atomic Power Station*

Recognising that self-reliance would take some time, it was decided to construct a BWR to start with, so that the experience of operating a nuclear power station in an Indian electrical grid could be gained. It was

only as an introduction to nuclear power that the Tarapur Station comprising of two 210 MWe reactors was built by a US firm on a turnkey basis. Since commissioning in 1969, these reactors have operated at an average capacity of over 50 per cent and have been supplying the cheapest non-hydel electric power in the country. Tarapur's performance has been wholly satisfactory and comparable with other stations of similar vintage elsewhere in the world. Its radiation levels are no different to those prevailing at similar plants in the U.S. During its 15 years of operation, the availability of the plant has been around 70 per cent. The total down time has essentially been due to refuelling (20 per cent) and faults in equipment, grid, etc. (10 per cent). Though the Tarapur reactors were originally rated at 210 MWe each, it would be more appropriate to assess its present performance factors on the basis of a lower capacity. Due to fuel availability problems and downgrading of some equipment, the present capacity of these imported units should be taken at around 160 MWe each.

### *Rajasthan Atomic Power Station*

Simultaneously with the above reactors, work was undertaken on two 220 MWe, PHWRs at Rajasthan. Most of the equipment for the first unit was imported from Canada. The first unit was commissioned in 1972 and commenced commercial operations from December 1973. During its initial years of operation, apart from several teething problems, its performance was severely affected by the voltage and the frequency fluctuations in the grid. Of the 251 outages, the unit experienced till end 1981, nearly a quarter of these were due to grid problems. Consequently, despite achieving a creditable capacity factor of 64 per cent in 1979, its overall performance till 1981 was far below the target: it could achieve an overall capacity factor of only 31 per cent during this period. In 1982, a very difficult technical problem arose on one of the endshields, which necessitated the unit to be shut down. Since then, repair works of the kind never undertaken anywhere else in the world are in progress. Before one criticises the performance of this unit, it must be realised that its design was based on a first generation reactor which was yet to be commissioned in Canada itself. It also had to operate under irregular conditions arising from an unsteady grid. In this respect, the first unit at Rajasthan should basically be considered as an imported prototype reactor, and the lessons learnt from it have been invaluable on our road to self-sufficiency in the design and operation of nuclear power reactors.

The second unit at Rajasthan was to incorporate a significant indigenous content. Due to export restrictions placed by some countries, a much larger indigenous content became inevitable and this could be achieved only at the expense of a slippage in its time schedule. After commissioning in October 1980, the unit commenced commercial operation in April 1981. Upto June 1984, it has experienced 52 outages over half of which were during 1982. Around 56 per cent of the outages are attributable to problems in conventional systems, 16 per cent due to grid problems and 22 per cent due to problems in the nuclear systems which are essentially akin to conventional systems but are required to have a higher reliability. Its operation to date should be considered as satisfactory. During 1983-84 it achieved an average capacity factor of about 64 per cent with its power level being maintained around 200 MWe. In recent times, its performance has further improved. The unit operated at ever 77 per cent capacity factor during January and May and at a record capacity factor over 88 per cent during August.

#### *Madras, Narora and Kakrapar Atomic Power Projects*

As mentioned earlier, the commissioning of the first 235 MWe unit at Kalpakkam (near Madras) last year, marks the culmination of all our efforts in becoming self reliant in the field of nuclear power. Though basically the two reactors at Kalpakkam are similar to the units at Rajasthan, several significant changes to improve their performance and safety have been incorporated. The performance of the first unit since January 1984 when it was declared commercial has been most satisfactory and a record capacity factor of 93 per cent was achieved in July. The second unit is in its final stages of construction and is expected to be commissioned next year.

Four more similar units are under construction — two at Narora in U.P. and two at Kakrapar in Gujarat. Several new concepts like earthquake proof design, and scaling up of some major equipment so as to be useful for 500 MWe units, have been introduced in these reactors.

#### *Overall Progress and Current Performance of Stations*

The problems faced with regard to imports and the efforts at indigenisation have stretched the time schedules of these projects and also escalated their costs. The technology for the fabrication of major components like calandria, end-shields, steam generators and fuelling machines had to be absorbed and adapted by Indian manufacturers. It is a

source of satisfaction that the Indian Industry is now fully capable of supplying all these critical nuclear components. With the standardisation of the 235 MWe unit, the problems faced earlier in terms of extended schedule are not expected to recur. As regards operation of these power plants, their continuously improving performances have generated new confidence.

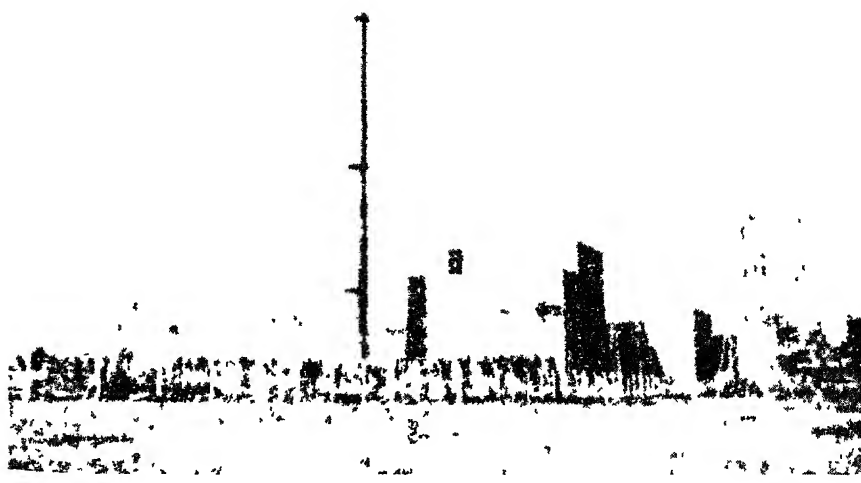


PLATE I View of the Tarapur atomic power station, Maharashtra

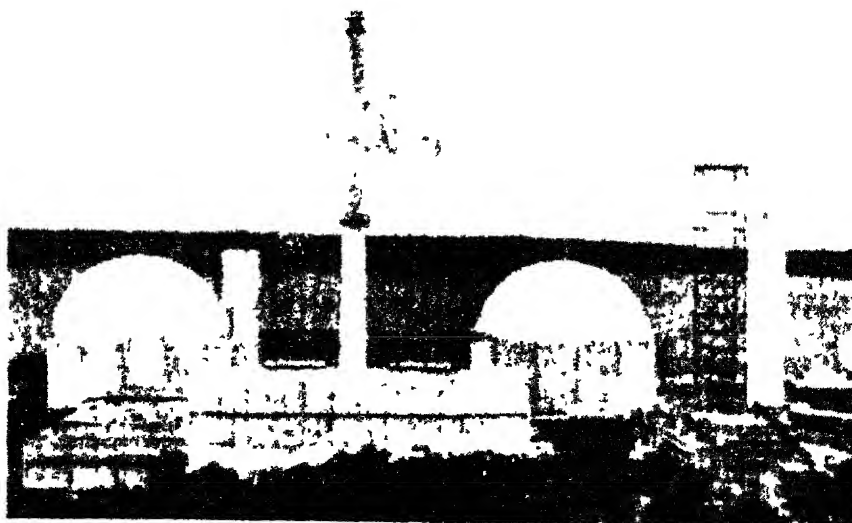


PLATE II View of the Rajasthan atomic power station at Kota, Rajasthan.

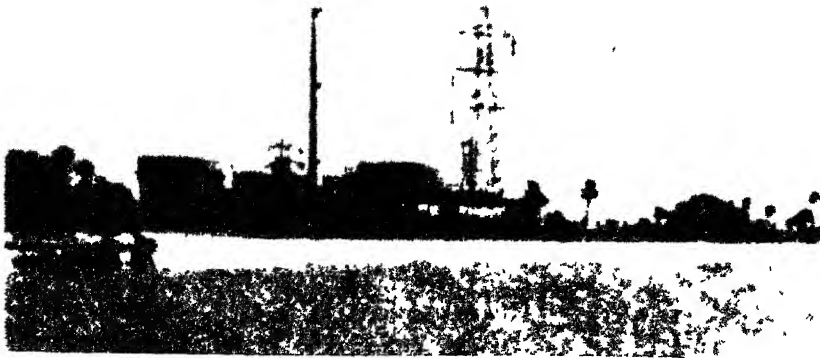


PLATE III View of the Madras atomic power project at Kalpakkam, Tamil Nadu



PLATE IV View of the Narora atomic power project under construction at Narora, Uttar Pradesh

The excellent performance of the operating nuclear power stations during this year (upto Sept. 1984) is attested by the capacity and availability factors achieved by them:

|                                | <i>Capacity factor</i> | <i>Availability factor</i> |
|--------------------------------|------------------------|----------------------------|
| Tarapur - 1                    | 54%                    | 92%                        |
| Tarapur - 2 (since refuelling) | 71%                    | 97%                        |
| Rajasthan - 1                  | 71%                    | 88%                        |
| Madras - 1                     | 67%                    | 79%                        |



PLATE V View of the heavy water plant at Baroda

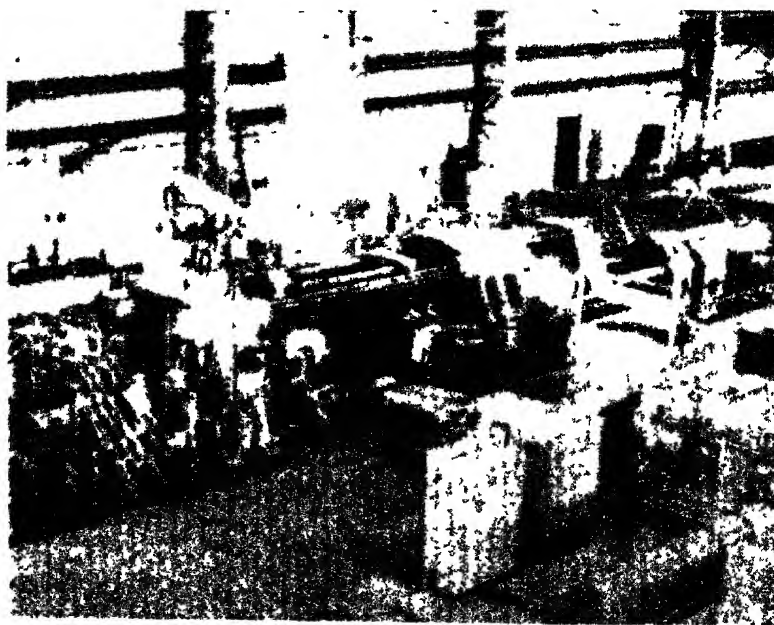


PLATE VI The 3150 tonne extrusion press of the zircaloy fabrication plant at the nuclear fuel complex, Hyderabad.

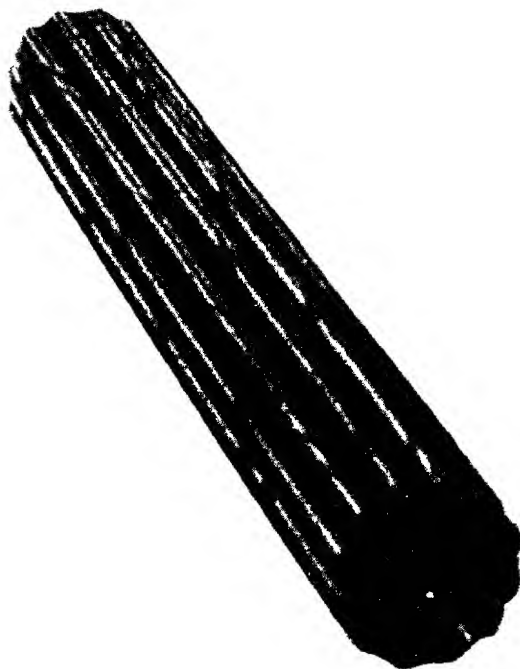


PLATE VII A natural uranium fuel bundle (for the pressurised heavy water reactors) manufactured by the nuclear fuel complex, Hyderabad.



PLATE VIII View of the Cirus, and Dhruva reactors at BARC, Bombay



Plate IX View of the fast breeder test reactor at the Reactor Research Centre, Kalpakkam, Tamil Nadu

It is a matter of pride that the indigenously constructed Madras I operated at a capacity factor of 93 per cent and an availability factor of 99.8 per cent during July 1984. The above figures can be contrasted with the all-India average figures for all operating power stations. The average capacity factor for 1983-84 was around 48 per cent and the average availability factor upto 1983-84 was about 69 per cent. The revenues earned by these nuclear power stations against the capital costs incurred on them are given in Table III.

**Table III**  
*Revenues earned by the nuclear power stations*

| Sl No | Station       | Operations commenced | Capital Cost incurred (Rs crores) | Approximate revenues earned till Sept 1984 (Rs crores) |
|-------|---------------|----------------------|-----------------------------------|--|
| 1     | Tarapur — 1&2 | 1969                 | 92                                | 327  |
| 2     | Rajasthan — 1 | 1972 (upto 1982)     | 73                                | 75   |
| 3     | Rajasthan — 2 | 1980                 | 92                                |  |
| 4     | Madras — 1    | 1983                 | 119                               | 32   |

(A) At present, sale of electricity from the operating nuclear power stations yields nearly Rs 14 crores every month.

(B) The revenues are proportional to the "rate per unit" at which the electricity boards buy electric power from the nuclear stations. Prevailing "rates" are of the order of



26 paise. 36 paise and 39 paise per unit for Tarapur, Rajasthan & Madras respectively.

- (C) The rates themselves are limited by statutory regulations.

It is seen that against a total capital investment of Rs. 376 crores on the Tarapur, Rajasthan and the first reactor at Madras, the total revenues earned till Sept. 1984 from these stations is nearly Rs. 534 crores. At present, the sale of power from these stations yields nearly Rs. 14 crores every month.

## SUPPORTING PROGRAMMES

### *Exploration, Mining and Processing of Uranium*

The exploration for uranium was initiated over three decades back, since the basic pre-requisite for any self-sufficient nuclear power programme is the availability of uranium. Initially, most of the exploration work of the Atomic Minerals Division was confined to the Singhbhum District of Bihar, but has now spread all over the country. During the last three decades, various ore reserves totalling to over 73,000 tonnes of  $U_3O_8$  in grades ranging from 0.015 per cent to 0.070 per cent of  $U_3O_8$  have been identified. These comprise of both 'inferred' and 'indicated' reserves though not all of them are commercially exploitable. The above reserves also include some by-product uranium recoverable from copper tailings. In addition, the phosphoric acid from phosphate based fertiliser plants, which contains 0.01 per cent to 0.03 per cent of  $U_3O_8$ , is also considered as a potential source.

Of the various ore reserves, about 49,000 tonnes of  $U_3O_8$  are considered commercially exploitable and are located mostly in the Singhbhum District of Bihar and in Madhya Pradesh, Meghalaya and Karnataka. The copper tailings from Roam-Rakha, Surda and Mosabani copper mines are expected to yield another 7800 tonnes of  $U_3O_8$ . Production of uranium from phosphoric acid could also yield about 1750 tonnes of  $U_3O_8$ .

The first mine and mill was set up at Jaduguda in 1967 by the Uranium Corporation of India Ltd. and is at present supplying  $U_3O_8$  needed for fabricating the fuel for the power reactors at Rajasthan and Kalpakkam. Work on another mine at Bhatin is in progress and plants to recover uranium from the copper tailings have been commissioned at Surda and Rakha. Despite the fact that the country is not endowed with

high grade uranium deposits, it has been possible to locate, mine and process even the low-grade ores on a commercial scale.

### *Fabrication of Fuel and Zircaloy Components*

Right from the early stages of the nuclear programme, indigenous capabilities were established in the crucial area of fuel fabrication. Today, the Nuclear Fuel Complex at Hyderabad supplies the fabricated fuel and zircaloy components required by all the operating nuclear power stations. Zircaloy is required to sheath the uranium oxide pellets before it is put inside the reactor. The present capacity of the plant is 90 tonnes of uranium fuel per year and is expected to be doubled by end-1985. The capacity for the fabrication of Zircaloy components is also being increased from 35 tonnes to 50 tonnes per year. The expertise acquired in this field is attested by the extremely high quality production that has been maintained with respect to fabrication of all nuclear fuel and zircaloy components. During 1983, there was not a single failure in any of the indigenous fuel bundles supplied to the RAPS-2 unit and the defect rate of all fuel bundles manufactured till May 1984 at NFC is just 0.10 per cent in contrast to 0.28 per cent encountered on the fuel bundles imported originally for the Rajasthan reactors.

### *Heavy Water Production*

The process of heavy water production is essentially similar to that of any chemical process, except that the design parameters include — high temperatures and pressures and highly corrosive or highly explosive fluids. Pilot plant studies in France and Germany had demonstrated the feasibility of the Ammonia-Hydrogen exchange process and as a number of ammonia-based fertilizer plants were being set up in the country during the 1970s, it was decided to construct scaled up versions of these pilot plants near the fertilizer units. While the heavy water plants at Baroda and Tuticorin adopted the French process, the Talcher plant was based on the German process. It must be emphasized that there was no alternative but to set up the heavy water plants based on pilot plant studies, as the know-how for the only commercially proven process at that time i.e. the Hydrogen Sulphide-Steam exchange process, was simply not accessible to us. Nonetheless, indigenous technology for even this process was later developed and a large heavy water plant based on this process has been set up at Kota. If one bears in mind that the plants at Baroda, Talcher and Tuticorin were the first commercial plants of that type in the world and that the Kota plant was the first one based on an indigenously developed

process, it was inevitable that this multi-pronged approach had to result in many teething troubles, some of which have been over publicised. Even now several factors remain outside the plants' control. For example, the identification and rectification of several design problems at the Talcher plant have been held up due to the acute power shortage in that area and the non availability of synthesis gas from the fertilizer plant to which the heavy water plant is linked. On the other hand, the performance of the plants at Baroda and Tuticorin during the past year has shown remarkable improvement. Their present performance justifies the confidence in setting up two more heavy water plants — one at Thal-Vaishet (with 110 Tonnes per year capacity) based on the Ammonia Exchange process and another at Manuguru (with 185 Tonnes per year capacity) based on the Hydrogen Sulphide exchange process. Today the country has sufficient expertise in various heavy water production technologies to design, construct and operate these plants satisfactorily without any foreign assistance.

### *Spent-Fuel Reprocessing*

Reprocessing and Waste Management are the last but vital stages of a nuclear fuel cycle and here again the technology is available only with a handful of countries. Years of sustained effort by our scientists and engineers has resulted in complete indigenous development of these operations as well.

Spent-fuel reprocessing provides the important link to the fast breeder power reactor programme and optimises the spent fuel storage capacity at various power reactor sites. An indigenously designed plant for reprocessing spent-fuel from research reactors was built at Trombay as early as 1964. This unit has recently been refurnished and expanded. Based on this experience, a large plant for reprocessing power reactor fuel was constructed at Tarapur and is at present reprocessing the spent fuel from the Rajasthan nuclear power station. Construction work on another 100 tonnes per year plant at Kalpakkam has commenced and it is expected to be operational around 1990.

### *Radioactive Waste Management*

The protection of the environment is an essential feature of the nuclear industry and it is in this connection radioactive waste management has received high priority from the very inception of our nuclear programme. Work was started in all three areas of waste management: low, intermediate and high-level radioactive wastes. For low and intermediate

level wastes, technology developed for their disposal is already in use at various nuclear facilities. All the stringent stipulations of environmental protections are being met at these locations. For the management of high-level radioactive waste, the concepts generally accepted the world over are:

- Immobilisation of the waste in appropriate matrices (vitrification),
- Interim on-site retrievable storage of solidified waste for about 25 years in a near-surface facility to allow dissipation of the decay heat, and
- Final disposal of solidified wastes in deep geological formations specifically selected for the purpose.

The first Waste Immobilisation Plant and the associated interim storage facilities are nearing completion at Tarapur. Future efforts will be directed to establishing an adequate degree of confidence in deep geological formations. The level of technology existing in all other countries is the same and final repositories for solidified high-level wastes are expected to be established only by the turn of the century. Though the complexity of managing high-level wastes is undoubtedly high, it should be kept in mind that its volume is very small. For example, the entire quantity of high-level waste resulting from one year's operation of a 4 x 235 MWe nuclear power station will occupy a volume of just 4 cubic metres.

### A 15-YEAR PROGRAMME FOR NUCLEAR POWER

From the experience gained in designing, constructing, and operating nuclear power stations, a stage has now been reached when the nuclear power generation capacity can be stepped up to meet the growing demand for electricity in the country. Towards this end, a detailed study was undertaken last year taking into account the following factors:

- the total electrical generation capacity required by the country by the year 2000,
- the quantum of uranium resources that have been identified for exploitation and mining,
- the capability and capacity of the indigenous industry,
- the experience gathered to date in the entire nuclear fuel cycle, and
- the economics of nuclear power.

This study indicates that a total installed capacity of 10,000 MWe by the year 2000 is feasible, subject to some pre-requisites. The proposal based on this study covers in detail every aspect of the programme such as the detailed physical plans in each sector, involvement of industry, manpower planning, financial outlays and finally, the infrastructure necessary to ensure its speedy implementation.

## PLANS IN EACH SECTOR

### *Nuclear Power Stations*

Nuclear power generation capacity at present is about 1100 MWe and projects under construction will increase this to 2300 MWe by the year 1990-91. The 15-year programme calls for setting up 12 more reactors of the standardised 235 MWe type and 10 reactors of 500 MWe capacity. The designs for the latter will be ready by 1987 and first 500 MWe reactor can be expected to go into operation by 1995.

### *Exploration, Mining and Processing of Uranium*

As mentioned earlier, the commercially exploitable uranium reserves are about 49,000 tonnes of  $U_3O_8$  and a further potential of 7800 tonnes of  $U_3O_8$  is available from copper tailings. After accounting for various losses during mining, milling, processing and fuel fabrication, these reserves are expected to yield about 38,000 tonnes of  $U_3O_8$ . A 10,000 MWe nuclear power programme will require about 37,000 tonnes of  $U_3O_8$  over the 25 years life time of the reactors. As such, the availability of uranium within the country will not be a constraint. The exploration activities will, however, have to be intensified to locate higher grade ores, and to identify larger deposits which will enable the minimisation of mining costs.

New Mines and Mills are being planned at Narwapahar and Turamdih (East) together with a full-scale recovery plant at Mosabani for the recovery of uranium from copper tailings. A total mining and processing capacity of 1800 tonnes of  $U_3O_8$  will be established by the end of the century, which will be sufficient to meet the fuelling requirements of all the nuclear power stations that will be operational by then.

### *Fabrication of Fuel and Zircaloy Components*

Fabricated fuel requirements for a 235 MWe PHWR are 61 tonnes of uranium in the first year and 33 tonnes of uranium per year subsequently, and those for the 500 MWe reactor, 115 tonnes and 61 tonnes respectively. Zircaloy requirements are 19 tonnes and 43 tonnes for the 235 MWe and

500 MWe PHWRs respectively. To meet the requirements for the 10,000 MWe programme, apart from the present expansion projects in progress, it is planned to expand them further so that a total capacity of 1500 tonnes of fabricated uranium fuel per year and 250 tonnes of zircaloy per year is available.

### *Heavy Water Production*

The initial heavy water inventories for the typical 235 MWe and 500 MWe reactors are about 255 tonnes and 485 tonnes respectively. These reactors require a make-up of around 8 tonnes per year and 15 tonnes per year respectively during their operation. The total estimated requirement of heavy water for the entire programme is about 13,000 tonnes and to meet this requirement, additional heavy water plants are planned to increase the total annual production capacity to about 1530 tonnes per year.

### *Spent-Fuel Reprocessing and Radio-active Waste Management*

In addition to the 100 tonnes per year spent fuel reprocessing plant under construction at Kalpakkam, it is planned to set up two additional plants near other power reactor locations, each with a nominal capacity of 400 tonnes per year. The first of these is expected to be operational by the year 1995 and the second by 2000.

In the field of waste management, in addition to the waste immobilisation plant at Tarapur, three additional plants with associated interim storage facilities are being planned. Two of these will be required to be operational by the year 2000. The final repository for solidified high-level radio-active wastes in a deep geological formation will need to be established only by the turn of the century and studies are already in progress to find suitable sites.

## INVOLVEMENT OF INDIAN INDUSTRY

A programme of this magnitude cannot succeed without a total commitment by the Indian industry. Subsequent to a dialogue held with their representative, they have unanimously supported and expressed their confidence in the practical feasibility of undertaking such a large programme. They have also highlighted the need for a firm long term commitment to such a programme so that their long-term financial viability is not adversely affected. The continuity of orders for the major components is a pre-requisite to ensure timely deliveries and reasonable

returns on their investments. It might be necessary to consider awarding cost-plus basis contracts for manufacture of components for the first 500 MWe reactor unit, which involve substantial developmental work.

### MAN-POWER PLANNING

The availability of specialised scientific and technical manpower has been found to be a major constraint in the undertaking of a large nuclear power programme. The programme envisages a total requirement of over 6,000 scientific and 29,000 technical personnel. Their recruitment, training, career development, retention and motivation have all been considered. However, given the quality of people available within the country and with the experience of training programmes at the Bhabha Atomic Research Centre (BARC), the recruitment of qualified personnel is not expected to pose major problems.

### FINANCIAL OUTLAYS AND REVENUES

To enable an overall assessment of the entire programme, the estimated levels of capital, revenues and O&M expenditure have been examined in great depth. The following details cover the plants and facilities under operation, those under construction and those that are planned to be taken up.

#### *Capital Expenditure*

The capital expenditure for the entire programme has been estimated at about Rs. 14,000 crores (1983 prices). The periodwise and sectorwise break-up of this expenditure are given in Figs. 2 and 3. It will be seen that the maximum amount of funds (about Rs. 7400 crores) will be required during the 8th plan period. (1990-95). Sectorwise the nuclear power reactors and heavy water production will account for nearly 85 per cent of the total expenditure.

#### *Operations Maintenance Expenditure*

The O&M expenditure in each of the sectors till the year 2000, total to about Rs. 8300 crores (1983 prices). The periodwise and sectorwise breakup are also shown in Figs. 2 and 3. As can be expected, the major expenditure under this head will commence only from the 9th plan period (1995-2000) onwards by which time many reactors will have become operational. The uranium exploration, mining and processing sector, the

fuel fabrication sector and the heavy water production sector will consume about 84 per cent of the O & M expenditure during the next 15 years.

### Estimated Revenues

The revenues from this programme will be from the sale of electric power. At present the cost of electricity (per unit) from Tarapur, Rajasthan and Madras Nuclear Power Stations are around 26 paise, 36 paise and 39 paise respectively. The electricity cost from Narora, is estimated at 50 paise and that for all future stations at 65 paise (Table IV). The total revenues for the entire programme have been estimated on these electricity costs. The costs of reprocessing and waste management have not been included, nor has any credit been taken for the plutonium that will be available from the spent fuel.

The estimated cumulative capital and O & M expenditure and the gross revenues from sale of electricity till the turn of the century are, shown graphically in Fig. 4. It will be seen that by the year 2000 the revenues from the sale of electricity will cross Rs. 14,500 crores thereby recovering the capital expenditure that would have been incurred till then. It is estimated that by the year 2025, a *net* revenue of over Rs. 60,000 crores (1983 prices) would have been earned by the programme.

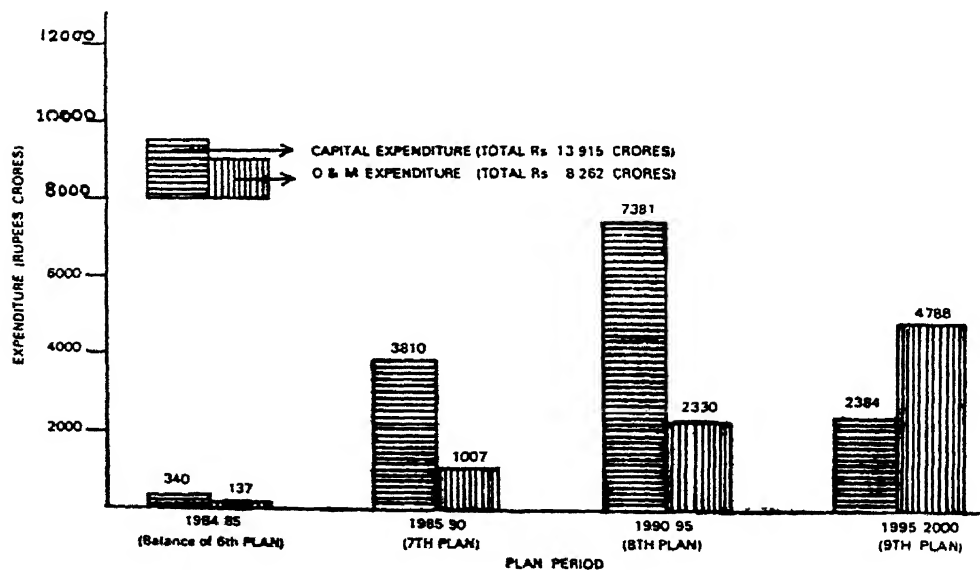


FIG 2 Projected capital and O & M Expenditure for the 10,000 MWe nuclear power programme periodwise break-up upto the year 2000



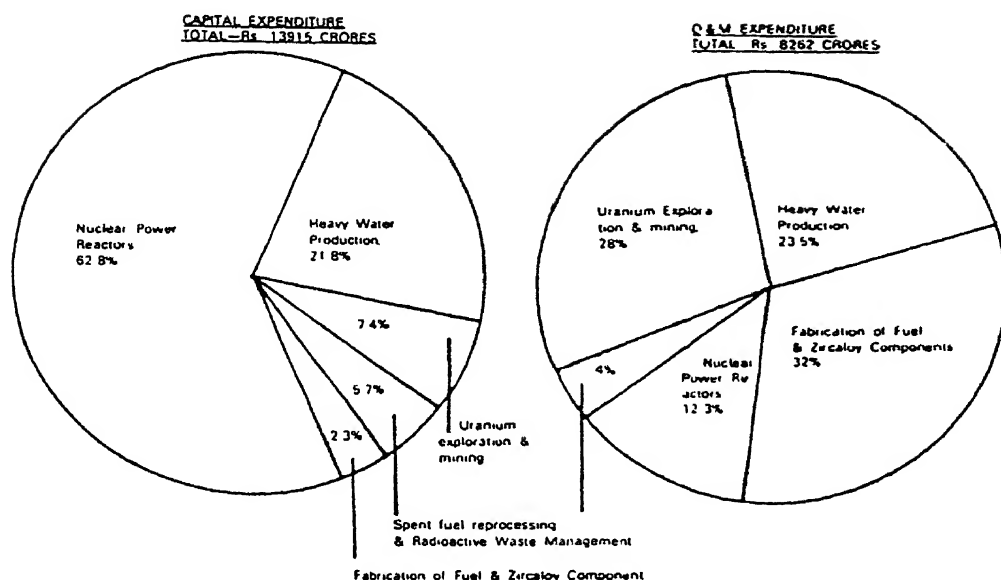


Fig 3 Capital and O & M expenditure for the 10,000 MWe nuclear power programme sectorwise break-up upto the year 2000

Table IV

*Cost of electricity from a  $2 \times 235$  MWe nuclear power station commissioned during the 1990's*

| Basis   |                                      |
|---|--------------------------------------|
| Project cost  | Rs 530 crores (ie) Rs 11,300 per KWe |
| Heavy water cost  | Rs 6635 per Kg.                      |
| Uranium fuel cost                                       | Rs 4545 per Kg.                      |
| Project construction time                               | 8 Years                              |
| Economic operating life                                 | 25 years                             |
| Energy sales  | 2780 million units per year          |
| Interest during construction                            | 6.9% per annum                       |
| Heavy water lease charges                               | 8% per annum                         |
| Return on capital employed                              | 12% per annum                        |
| Rate of depreciation                                    | 3.6% per annum                       |
| Methodology   | Return on investment method          |
| <b>Break up of cost of electricity (paise per kwhr)</b> |                                      |
| Return on capital                                       | 28                                   |
| Heavy water lease charges                               | 10                                   |
| Depreciation  | 8                                    |
| Decommissioning expenses                                | 1                                    |
| Fuel Consumption  | 11                                   |
| Heavy water consumption                                 | 4                                    |
| Operation and Maintenance                               | 3                                    |
| Total   | 65 paise per kwhr                    |

Fixed charges=47 paise

Operating expenses=18 paise kwhr

## PRE-REQUISITES FOR SUCCESSFUL IMPLEMENTATION OF THE PROGRAMME

### FACTORS RESPONSIBLE FOR SLIPPAGES IN THE PAST

In the light of past experience, the primary objective in such a programme will be the necessity to adhere to project schedules. Fears expressed on many occasions that any increase in costs of heavy water or the uranium fuel will result in prohibitively high cost of electricity are not valid as nuclear stations being capital intensive, are more sensitive to project schedules. Experience has shown that factors which affect the schedules are:

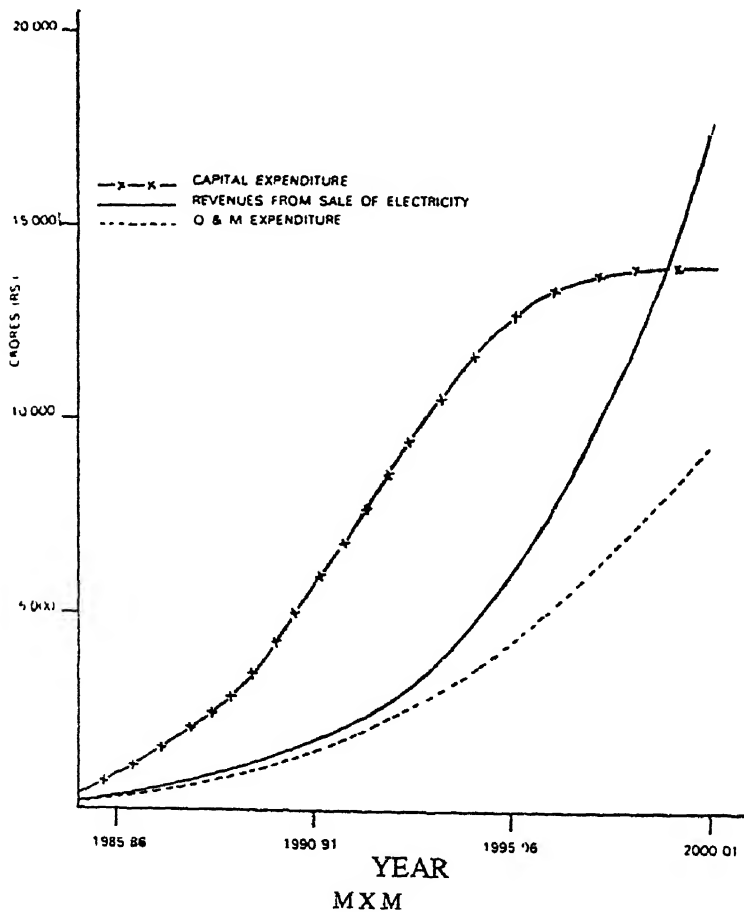


FIG. 4 Projected capital expenditure, O&M expenditure and revenues from sale of electricity (all cumulative in crores rupees) for the 10,000 MWe nuclear power programme.

- Delay in selection and acquisition of sites,
- Delays in deliveries of major equipment. These have primarily been due to the manufacturer finding the job unprofitable and therefore losing all motivation in completing it. Another reason has been

shortage of certain raw materials that ultimately need to be imported,

- Need for standardising designs,
- Problems in recruitment, retention and motivation of qualified personnel, and
- Procedural delays inherent in the organisation

### THE PRE-REQUISITES

The proposed programme addresses itself to all these problems and specific suggestions have been made against each of them. Specifically, the following steps have been identified as being essential to the programme's success:

#### *Long Term Commitment*

Apart from the point made earlier that power plans have to be drawn up on a long time frame, only a firm commitment will convince the industry of a "steady and predictable" demand for the major equipment. Fabricating 'one-off' involves long fabrication times as the manufacturer has to gear up his facilities to meet the technical specifications. The manufacturer will be reluctant to invest on costly equipment, unless there is an assured long term demand for such equipment. Consequently, the two basic factors, i.e., production planning for timely deliveries and economics of production, dictate that a *long term commitment is a prime necessity*. Piecemeal undertaking of nuclear projects will lead to long time delays and cost overruns.

Some *raw material*, which constitute a very small percentage, have to be imported. As problems have been faced in the past in this area, it would be necessary to import the total quantity under a blanket licence and store the same for future use.

#### *Selection of Sites and Advance Land Acquisition*

Approval of the *selection and of all the sites* for power reactors, heavy water plant, fuel fabrication facilities, etc. will be required at the initial stage itself, though the actual site developmental work at each location will commence as per the schedule. This is necessary to achieve the construction schedule of 7 to 8 years for power reactors, 5 years for heavy water plants, 5 to 9 years for uranium mines and mills and about 8 years for spent fuel reprocessing and waste management plants.

#### *Standardising Designs*

As of date, the 235 MWe design adopted for the Narora and Kakrapar projects has been *standardised* for all future 235 MWe reactors. In the 500

MWe designs, the features incorporated in the Narora and Kakrapar reactors will enable its early standardisation, once the prototype is developed. It is expected that the design specification for this unit will be ready by 1987 for the purposes of tendering and placing orders.

### *Man Power Development*

To attract, and thereafter retain 6000 scientific and 29,000 technical personnel would be difficult in view of competing attraction from abroad and the private sector. Fortunately the manpower loss amongst scientists and engineers has been small—of the order of 2 to 5 per cent. Efforts are being made and will need to be continued to be made to improve the service conditions, prospects for career development and amenities like housing, transport etc.

### *Organisational Restructuring*

To streamline the decision making process, it will be necessary to effect a certain amount of decentralisation by way of adequate delegation of authority and identify responsibility and accountability at various levels. It will expedite the process of site selection, design and engineering, procurement, construction, erection/commissioning and operation of the various projects. As a first step, the erstwhile Power Projects Engineering Division has been restructured into a Nuclear Power Board. Sufficient powers have been delegated to carry out these tasks

## **SOME RELEVANT ISSUES IN THE MIND OF THE PUBLIC**

The proposal to invest on such a large scale in the background of the country's limited financial resources will obviously raise some relevant issues in the mind of the public which need to be examined. First is the question whether nuclear power is safe and the second is whether the projected schedules can be achieved.

## **SAFETY ASPECTS OF NUCLEAR POWER**

### *'Imaginable' and 'Expected' Accidents*

No other safety issue has been so thoroughly scrutinised during the past few years as the one associated with nuclear power. The correct perspective has somehow been side tracked in this process and lingering doubts prevail in the mind of the public. Most discussions on safety of nuclear power have been focussed on the possibility of catastrophic nuclear accidents which are "imaginable" but extremely unlikely. If such

"Imaginable" arguments were to be applied to other fields like Aviation, Rail & Road Transport, Mining etc.—we would have to stop all these activities. One would therefore have to distinguish between "expected accidents" and "imaginable accidents".

Expected accidents are those like the crash of an aircraft, collision of trains, mining disasters, all of which have occurred but have not stopped any of these activities. The highly publicised "Three Mile Island" event was one such event, which in the final analysis was an economic catastrophe and did not affect the safety of the public. All nuclear power stations are designed in such a manner that "expected" accidents would have no consequences on the public, designed as they are on the fail safe principle.

On the other hand there can be no limit to "imaginable" accidents the limit being only the human imagination: One can for example, imagine that an airbus with full fuel tanks and carrying over 300 passengers which has just taken off from the Palam Airport crashing into a fully packed cricket stadium. The consequences of such an event would be catastrophic, but most of us would dismiss such a possibility as having an extremely low probability. Similar events that are so often quoted in various discussions on nuclear safety pertain to such 'imaginable accidents'. These have no rational basis whatsoever—but it does make exciting fiction, and unfortunately makes the public skeptical about nuclear power. Sometimes one cannot help feeling that such 'imaginable accidents' are propagated more for political reasons. To go by simple facts and its record to date, nuclear industry all over the world *has had the safest record* amongst all industrial activities. The present generation of nuclear reactors are designed to have a negative temperature coefficient of reactivity. If due to any reason, the temperature of the reactor core increases, the inherent characteristics of the system will automatically bring down its power level. The safety record of nuclear reactors is thus largely attributable to such in-built safety features.

### *Radiation Levels*

As regards the effects of 'Radiation' itself which is a related subject, one should first be aware that mankind has probably gathered as much if not more knowledge on this subject as say on any other harmful agent. A powerful network of both international and national organisations is supervising and controlling this aspect. The International Commission on Radiological Protection (ICRP) recommends the limits for indicative

exposure. The International Atomic Energy Agency (IAEA) in consultation with the World Health Organisation (WHO), International Labour Organisation (ILO) and other world bodies, prepares Basic Safety Standards related to the ICRP recommendations. These Basic Safety Standards are adopted by the national organisations like the recently constituted Atomic Energy Regulatory Board (AERB). In fact few processes in the world are monitored and controlled so vigorously both at national and international levels. Yet, despite this, 'Radiation' continues to arouse fears in the mind of the public. It is probably due to the fact that it can neither be seen nor felt like other hazards.

But on the other hand, it is a fact that *all of us* are exposed to radiation in some form or other during our day-to-day life—the magnitude varying from place to place : Elements present naturally in our food, water and air expose us to nearly 25 millirems per year and cosmic radiation also exposes us to nearly double that amount. Anyone travelling by air gets exposed to a higher radiation in comparison to others journeying on land. Inhabitants of concrete or brick houses are exposed to a higher radiation compared to those living in wooden houses. Milk contains 200 more times radioactivity than drinking water and three times as much as in beer. One of the largest man-made sources which exposes modern man to radiation is that resulting from Medical procedures like Diagnostic X-rays, Radiotherapy etc. The inhabitants of the 55 km strip of beach in Kerala, which contains the worlds' richest deposits of thorium, receive about 400 millirems per year compared to the 500 millirems per year receivable by most personnel at a nuclear facility. In fact, mankind has been exposed to radiation more due to nuclear weapons testing rather than from nuclear power stations. For example, radiation dosage from the nuclear weapons testing in 1961-62 (when it was its peak) equalled that which would have resulted from nearly 350 reactors of 235 MWe each operating for 1000 years. If one compares the risk of being exposed to the permissible radiation levels with other risks, there are many other causes in our day-to-day life which are more hazardous. Some of these are shown in Table - V.

**Table V**  
*Comparison of risk from radiation with risks due to other activities*

| Cause   | Chances of death/year |         |
|---|-----------------------|---------|
| Smoking 20 cigarettes/day                               | 1 in                  | 200     |
| Accidents in deep sea fishing                           | 1 in                  | 400     |
| Natural causes, 40 year old                             | 1 in                  | 500     |
| Accidents on the road                                   | 1 in                  | 5,000   |
| Accidents in the home                                   | 1 in                  | 10,000  |
| Accidents at work                                       | 1 in                  | 20,000  |
| Radiation work at nuclear facilities (0.5 rem per year) | 1 in                  | 20, 000 |

Source Adapted from the document "Living with Radiation", published by the National Radiological Protection Board, U K

These comparisons and statistics are given only to point out the extreme caution with which the permissible levels have been set. The actual emissions in practice are in fact lower. The radiation emission from a nuclear power station is comparable to that from a coal fired station—but in both cases, the actual dose received is less than the variations in the annual dose received from all natural sources. The only difference is that a malfunction in a nuclear power station could increase this emission, whereas a similar event in the coal fired station would not. It is for this reason, safety precautions are almost an obsession in nuclear reactors and all possible steps are taken to observe it. Experience with hundreds of reactors now operating all over the world has shown that the nuclear reactor is a safe and clean way of providing electricity.

### CONFIDENCE IN IMPLEMENTING THE PROGRAMME

Doubts on the confidence in implementing our programme arise as a consequence of slippages and cost overruns encountered in the past on various projects. The delays were inevitable in view of the scarcity of financial resources combined with the multipronged approach to achieve self-reliance in all aspects of nuclear power. Now that the country has achieved this primary goal, there is no reason why this performance cannot be repeated—but much faster. Cost escalations are essentially the result of project slippages and this is borne out by experience elsewhere. Table 6 and 7 give the cost escalation and project slippages on nuclear power projects in India compared with similar phenomena reported on

some projects in USA. It will be seen that the uncertainty in a firm long term programme has resulted in severe slippages in the USA thereby resulting in tremendous escalation in project costs. The figures in these tables speak for themselves: while this is neither an explanation nor a justification, it must be realised that this phenomenon is not confined to this country alone.

On the other hand, one could argue that nuclear power programme in France and Japan have been highly successful and implemented on schedule: at one stage, France was commissioning a reactor every six months. The reasons are not hard to find: a national consensus on the need for nuclear power and a firm commitment by the Government to the industry were the primary factors that contributed to their successes.

**Table VI**

*Comparison of escalation of nuclear power project costs in India with recently reported increases in the cost of some nuclear power stations in USA*

| Sl No.            | Project         | Originally estimated cost (Rs. crores) | Final or present estimated cost (Rs. crores) | Percentage escalation |
|-------------------|-----------------|--|--|-----------------------|
| 1.                | Tarapur-1 & 2   | 53                                     | 92   | 73%                   |
| 2.                | Rajasthan-1 & 2 | 100                                    | 165  | 65%                   |
| 3.                | Madras-1 & 2    | 132                                    | 246  | 86%                   |
| 4.                | Narora-1 & 2    | 210                                    | 400  | 90%                   |
| 5.                | Kakrapar-1 & 2  | 383                                    | 383  | Nil                   |
| <i>In U S A**</i> |                 |  |  |                       |
| 1.                | Shoreham        | 241                                    | 4000   | 1560%                 |
| 2.                | Midland         | 267                                    | 4400   | 1548%                 |
| 3.                | Vogtle          | 660                                    | 6600   | 900%                  |
| 4.                | Diablo Canyon   | 450                                    | 4400   | 878%                  |
| 5.                | Seabrook        | 973                                    | 5800   | 496%                  |
| 6.                | St. Lucie       | 360                                    | 1400   | 289%                  |
| 7.                | Palo Verde      | 2800                                   | 6000   | 114%                  |

\*\* According to a survey of 47 Nuclear Power Stations in the U.S.A., it is reported that 36 of them costed at least 100% or more than originally estimated and 13 of these costed over 300% more than the original estimates. It is also reported that the U.S. Industry considers those at St. Lucie and Palo Verde as examples of successful limitation of cost increases under present conditions.



## PERSPECTIVE PLANS FOR NUCLEAR POWER AND SUPPORTING R & D ACTIVITIES

### PERSPECTIVE PLAN – THE FAST BREEDER REACTOR

It must be recognised that a nuclear power programme based solely on the natural uranium fuelled PHWR cycle cannot be sustained indefinitely to meet the country's growing electricity requirements into the next century. In the absence of commercially viable alternative energy sources for this purpose (apart from the limited hydro and coal potentials), nuclear power will have to play an even greater role in meeting these requirements. Towards this end, developmental work on the next phase of the nuclear power cycle is already in an advanced stage.

**Table VII**

#### *Slippages in schedules of Nuclear Power Projects*

The cost escalations reported in Table VI are basically a consequence of slippage in project schedules This Table gives the extent of these slippages

#### *Experience in India*

| S. No. | Project       | Work commenced in | Scheduled completion by | Actual/Expected completion in | Extent of Slippage |
|--------|---------------|-------------------|-------------------------|-------------------------------|--------------------|
| 1.     | Tarapur – 1   | 1964              | 1969                    | 1969                          | Negligible         |
| 2.     | Tarapur – 2   |                   |                         |                               |                    |
| 3.     | Rajasthan – 1 | 1964              | 1969                    | 1972                          | 3 years            |
| 4.     | Rajasthan – 2 | 1967              | 1973                    | 1980                          | 7 years            |
| 5.     | Madras – 1    | 1967              | 1973                    | 1983                          | 10 years           |
| 6.     | Madras – 2    | 1971              | 1976                    | 1985                          | 9 years            |
| 7.     | Narora – 1    | 1974              | 1981                    | 1987                          | 6 years            |
| 8.     | Narora – 2    | 1974              | 1982                    | 1988                          | 6 years            |
| 9.     | Kakrapar – 1  | 1981              | 1990                    | 1990                          | Negligible         |
| 10.    | Kakrapar – 2  | 1982              | 1991                    | 1991                          | Negligible         |

Consequent to the standardisation of design and the acquisition of manufacturing know how by Indian manufacturers, future power projects are expected to take 7/8 years.

#### *Experience reported in the U.S.A., France and Japan*

- In the U.S.A
- Most projects listed in Table VI are reported to have slipped by 6 to 16 years.
  - Under present circumstances it is reported that nuclear power stations are estimated to take nearly 10 years for completion

- In France & Japan
- The schedules and costs are generally being maintained.
  - Nuclear Power Projects are reported to be completed in 5/7 years.
  - This has essentially been attributed to the drawing up of and commitment to, firm long term programmes.

The strategy envisaged for this is a simple one. Reprocessing of the spent uranium oxide fuel from the operating PHWRs will yield plutonium which itself can be used to fuel another type of reactor called the Fast Breeder reactor (FBR). In the FBR, a mixture of Plutonium and the unused uranium-238 from the PHWRs will be used as fuel for generating electricity. The FBRs will also have a blanket of Thorium oxide which when exposed to radiation inside the reactor, will finally yield Uranium-233 which in turn can be used to fuel a second generation of FBRs. The 10,000 MWe programme will yield enough plutonium to set up a 1000 MWe FBR every year. The present uranium reserves in the country can support nearly 3,50,000 MWe of the first generation FBRs in the next century. As regards the second generation of FBRs, using the U-233 extracted from the exposed Thorium Oxide, the abundant Thorium reserves of the country will probably cater to the country's entire demand of electrical energy by the end of the next century.

#### DEVELOPMENT WORK ON FAST BREEDER REACTOR TECHNOLOGY

As a step towards the second phase of the nuclear power programme, a separate Reactor Research Centre (RRC) was established at Kalpakkam for undertaking all the necessary development work connected with the fast breeder programme. Several laboratories for physics, metallurgy, reactor safety, sodium technology, reprocessing, radiochemistry, radiometallurgy and other related fields such as safety have already been set up at RRC, for carrying out research and development. A 15 MWe fast Breeder Test Reactor (FBTR) at this centre is expected to attain criticality by the end of this year.

The Plutonium Metallurgy Laboratory at BARC has undertaken the complex operation of fabricating the fuel requirements for the FBTR. A new mixed carbide fuel containing 70 per cent plutonium, 30 per cent uranium has been indigenously developed at this facility. India will be the only country in the world to use carbide fuel in the fast breeder reactors. Thorium oxide fuel elements for being used as 'Blanket' have already been fabricated at another facility in BARC. Experimentation in this area

will lead to the proving of the above stated strategy for utilisation of the abundant Thorium reserves in the country. If it can be demonstrated that the FBRs are safe and economical, India will indeed have solved her energy problems for a long time to come.

### CONSTRUCTION OF A PROTOTYPE 500 MWE FAST BREEDER POWER REACTOR

Experience gained at the RRC has provided the confidence to undertake the design and construction of a large prototype fast breeder power reactor. The feasibility report for such a reactor of 500 MWe capacity has recently been prepared. This report also outlines plans for taking up all the supporting developmental activities like fuel fabrication, spent fuel reprocessing and fabrication of nuclear components. It is expected that first Fast Breeder Power Reactor will come into operation by the end of this century.

### SUPPORTING R & D ACTIVITIES

Whilst the RRC will provide the R & D support for the fast breeder reactor project BARC will continue to provide R & D support for the PHWR programme. Even now a large portion of the R & D activities at BARC are devoted to the heavy water nuclear fuel cycle, starting from uranium recovery to waste management. The operational problems encountered in the various operating plants are taken up for analysis and solution by the specific Divisions at BARC. Without this supporting R & D activity, continued trouble free operations of many of the nuclear plants would not be possible. Experience has shown that R & D is needed not only during the design and development stages, but requires to be continued even during the operation of any facility.

### CONCLUSION

In this lecture we have discussed the present status of nuclear energy with respect to its economics and safety. There is no doubt that when the Fast Breeder Reactor goes into operation towards the end of the century, nuclear energy will have fulfilled its role in providing a source of cheap and safe electric power. However, we cannot rest on these developments alone. It is quite likely that in due course of time even the existing non-renewable energy resources will get exhausted. As has happened in science, there is always the possibility of an even better source of energy than fission being discovered. Such a possibility is fusion energy. It is for

this reason we have started the programme on the study of fusion in various parts of the country, particularly so at the Centre of Advanced Technology at Indore. It is not the scope of this lecture to comment on whether fusion will become an economically viable source of electric power now or even in the early part of the next century. It has not yet been demonstrated that controlled fusion can be achieved. We have to continue our work in this field with greater emphasis since its potential is tremendous.

It is now over 40 years since the power of atom has been harnessed for the benefit of mankind. And yet, even after having demonstrated that it is a safe, clean and economic way of producing electric power, it is difficult for anybody who has to deal with scientific planning to believe that nuclear energy is not an answer to our present day energy problems. The opposition to nuclear energy in our country does not seem to stem from a proper study of safety or economics. It seems to arise more from personal prejudices and biases rather than due to any rational and balanced analysis. It is a sad commentary on the quality of public discussion in this country that arm-chair experts with little practical experience or knowledge offer advice on topics ranging from how to run a power station or how to launch a satellite or for that matter how to cultivate land.

Contrary to the general impression that the Department of Atomic Energy is a secretive one, the reality is that it is practically as open as any other government department in this country. Our Annual Reports, Performance Budgets, Reports of the Estimates Committee and answers in the Parliament are all public documents available to anyone who cares to do a balanced and critical examination of our plans and performance. As a scientific body, such critical examination is not only welcome, but is desired by us. Unfortunately, there is a tendency amongst these experts to pick facts and figures out of context, usually from interested foreign publications. This has led to sweeping generalisations, thereby distorting the overall perspective.

What I have tried in this lecture is to convey the perspective that the demand for electricity in our country is exponentially growing and has to be met to keep up, if not improve, our economic growth rate. All technically and commercially viable forms of electricity generation should be explored by us, keeping in view the potentials and limitations of each of the alternatives. Exotic alternatives being widely discussed like Solar, Wind, Tidal etc. are appropriate only for localised applications requiring

small quantities of electricity. But they cannot cater to bulk electricity generation on a commercial scale. At the moment, nuclear energy is the only alternative which to a large extent will ensure the availability of cheap electricity in the years to come.

It is clear from my talk that developed countries are going ahead with exploitation of nuclear energy and there is definitely no sign of reduction in the growth of nuclear power in these countries. Countries which have succeeded in expanding nuclear power rapidly have already benefitted by it. It is against this background that our 15 year nuclear programme has to be considered. After all, 10,000 MWe of nuclear power by the year 2000 is a very modest proposal. Our concern is not that we have chalked out a nuclear programme which is too large for the country. On the contrary, one fears that even by the year 2000, shortage of electric power will continue despite having introduced nuclear power on a significant scale.

In today's world electric power is a prime requisite for any self-respecting country which desires to have a reasonable standard of living based on self-sufficiency. It is indeed a measure of prosperity of any society. Inevitably, nuclear power in one form or the other will be the only answer.

I would like to acknowledge the assistance of Shri K Muralidhar and Shri P K Bhatnagar in the preparation of this paper.



Sukh Dev (b. 17 June 1923) did Ph.D. (1950), from Panjab University; D.Sc. (1960), Indian Institute of Science, Bangalore. He is INSA-S.N. Bose Research Professor, Department of Chemistry, IIT, New Delhi (1988- ).

Sukh Dev discovered several new structural types in terpenoids; enunciated 'the absolute stereochemistry biogenetic rule'; determined the structure of lac resin; isolated and characterized kodo-millet toxins; characterized the hypolipaeamic compounds of *Commiphora mukul*, which led to the development of a modern drug; discovered several new reactions

with polyphosphoric acid; pioneered studies of organic reactions in a solid matrix; and introduced silver nitrate-silica gel for olefin chromatography. Authored/Co-authored 'Monoterpenoids' (2 vols), 'Diterpenoids' (4 vols), and 'Triterpenoids' (2 vols) (all by CRC Press);

Sukh Dev is Fellow, Intra-Science Research Foundation, USA (1971-76), Indian Academy of Sciences, Third World Academy of Sciences; Founder Fellow, Gujarat Science Academy; President, Indian Chemical Society (1978-79); has been member of INSA Council (1986-88). He is the recipient of Shanti Swarup Bhatnagar Prize (1964); Acharya P.C. Ray Memorial Medal (1970); Dr. K.G. Naik Gold Medal (1977); Vishwakarma Medal (INSA) (1979); Earnest Guenther Award (American Chemical Society) (1980); IISc Distinguished Alumni Award (1980); VASVIK Award (1980); FICCI Award (1980); Professor T.R. Seshadri 70th Birthday Commemoration Medal (INSA) (1982); Meghnad Saha Medal (INSA) (1987); Third World Academy of Sciences Award (1988); Srinivasa Ramanujan Birth Centenary Award (Indian Science Congress Association) (1992).

*Sukh Dev was elected to the fellowship of the Academy in 1968.*

**Address:** Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi 110016; Tel: (011) 6862208; Res: C-600, New Friends Colony, New Delhi; Tel: (011) 6847119.

## AYURVEDA AND MODERN DRUG DEVELOPMENT

SUKH DEV FNA

I am keenly aware of the honour done to me by the Academy in selecting me for the Meghnad Saha Medal for the year 1987. I feel all the more honoured, as it links me with the name of one of our most illustrious scientists, who was at once a great intellectual and a dedicated humanist.<sup>1</sup> I propose to speak to you to-day on *Ayurveda* and Modern Drug Development, a topic relevant to the current developmental needs of the country, and an area in which we have been interested for the past two decades or so.

Before proceeding further, it seems appropriate that I define the term modern drug, at least, in the context of the present lecture. The term drug has been defined in legal terms in various regulatory manuals,<sup>2</sup> but such definitions are not relevant for the present purpose. In the absence of a readily accessible (to me) suitable definition, I may define modern drug as a *clinically efficacious chemical entity or mixture, of synthetic or natural origin, administered as such or admixed with other entities or vehicles, and capable of being produced in a reproducible form under good analytical control*. The term *clinically efficacious* is meant to convey that the said product has undergone pharmacological, toxicological and clinical screening as per accepted parameters of modern medicine. Thus, a preparation based on a natural product and shown to be clinically efficacious and produced in a standardized form becomes a modern drug.

### ROLE OF NATURAL PRODUCTS IN MODERN MEDICINE

At present, it is generally believed by the lay public that most of the modern drugs are of synthetic origin and, that drugs of natural provenance are no longer important. This, in fact, is not true.<sup>3</sup> An analysis of prescriptions dispensed from community pharmacies in USA was carried out<sup>4</sup> in 1973, and this showed that as many as 41 per cent prescriptions contained one or more products of natural origin as the therapeutic agent. Of these prescriptions, 25 per cent were based on drugs from higher plants, some 13 per cent represented metabolites of microbes and, about 7 per cent were of animal origin. The total value of these prescriptions was estimated to be of the order of US \$ 5 billion (Rs. 6500 crores). Another later reference<sup>5</sup> mentions that 25 per cent of the 200 most frequently

prescribed drugs in the USA, were of natural origin. The situation would appear to be similar for many other countries including India, and as a matter of fact, this figure may even be higher for Soviet Union, Germany, Italy, China and Japan. Thus, natural products represent an important segment of modern drug armamentarium.

As already mentioned, drugs of natural provenance have been obtained from three different sources : higher plants, microbes and animals. Hormones, enzymes, antibiotics used in medicine to-day owe their origin to studies on animals and microbes. These contributions are some of the most important in modern medicine. However,, in the context of this lecture we are mostly concerned with drugs from plants. More than 75 pure compounds derived from higher plants find their place in modern medicine, and some of the traditional ones are shown in Table I. However, it may be noted that  $\Psi$ -ephedrine, ephedrine, emetine, caffeine, theobromine, theophylline and papaverine are now mostly produced synthetically for economic reasons. All compounds listed in Table I were introduced in modern medicine between 1850 and 1950. Beginning with 1950 science has made rapid, accelerating strides along all its facets and medicine has been no exception. Breakthroughs in a discipline often result from outstanding progress in related areas. Advances in physics, chemistry and biology, coupled with progress in isolation technology, analysis, characterisation and structure recognition have led to great advances in drug development. Thus the period 1950-70 saw the introduction of approximately one hundred basic new drugs in the U.S. market.<sup>6</sup> However, this list contained no more than five drugs (reserpine, deserpidine, rescinnamine, vinblastine and vincristine) derived from higher plants. Though, these compounds represent outstanding breakthroughs in modern medicine, the overall contribution from higher plants is clearly not impressive. This is attributed mainly to the fact that many pharmaceutical companies have been reluctant to invest in a major way, because of problems associated with plant collection, standardization and supply. However, the total effort, including investigations carried out in universities and research institutions not belonging to pharmaceutical companies, expended on biologically active compounds from plants has not been inconsiderable. A rather large number of new organic compounds exhibiting a variety of pharmacological activity have been isolated and characterized.<sup>5,7-9</sup> To emphasise the worthiness of these investigations, I have listed in Table II, types of biological/pharmacological activities (except anticancer) reported for compounds from higher plants, during the



three year period, 1977-79, covered in a review.<sup>6</sup> The economic importance of such studies can be gauged from the fact that in 1975 alone over 400 patents were issued for substances isolated from plants alone.<sup>4</sup>

**Table I**  
*Important traditional active plant principles<sup>a</sup>*

| No. | Drug                      | Plant source                |
|-----|---------------------------|-----------------------------|
| 1.  | Codeine                   | <i>Papaver somniferum</i>   |
| 2.  | Atropine                  | <i>Hyoscyamus muticus</i>   |
| 3.  | Ψ-Ephedrine <sup>b</sup>  | <i>Ephedra spp.</i>         |
| 4.  | Ephedrine <sup>b</sup>    | <i>Ephedra spp</i>          |
| 5.  | Hyoscyamine               | <i>Hyoscyamus muticus</i>   |
| 6.  | Digoxin                   | <i>Digitalis lantana</i>    |
| 7.  | Hyoscine                  | <i>Datura metel</i>         |
| 8.  | Digitoxin                 | <i>Digitalis purpurea</i>   |
| 9.  | Pilocarpine               | <i>Pilocarpus jaborandi</i> |
| 10. | Quinidine                 | <i>Cinchona spp</i>         |
| 11. | Quinine                   | <i>Cinchona spp</i>         |
| 12. | Emetine <sup>b</sup>      | <i>Cephaelis spp</i>        |
| 13. | Caffeine <sup>b</sup>     | <i>Thea sinensis</i>        |
| 14. | Theobromine <sup>b</sup>  | <i>Theobromo cacao</i>      |
| 15. | Theophylline <sup>b</sup> | <i>Coffea arabica</i>       |
| 16. | Papaverine <sup>b</sup>   | <i>Papaver spp</i>          |
| 17. | Colchicine                | <i>Colchicum autumnale</i>  |

<sup>a</sup> Based on ref 4

<sup>b</sup> Now mostly produced synthetically

**Table II**  
*Biological activities encountered (1977-79) during screening of higher plants for active compounds*

| No. | Biological Activity | Number of Biologically Active Compounds Isolated |
|-----|---------------------|--|
| 1   | Antibacterial       | 17   |
| 2   | Antifertility       | 2  |
| 3   | Antifungal          | 14   |
| 4   | Antiinflammatory    | 7  |
| 5   | Antiviral           | 3  |
| 6   | Antiulcer           | 3  |
| 7   | Cardiovascular      | 5  |
| 8   | CNS Depressant      | 5  |
| 9   | Hypoglycemic        | 1  |
| 10  | Hypotensive         | 13   |
| 11  | Hypocholesterolemic | 4  |
| 12  | Immunosuppressive   | 1  |
| 13  | Spasmolytic         | 3  |

Quite often, biologically active natural products have served as models for improved drugs. Corticosteroids<sup>10</sup> and semisynthetic antibiotics<sup>11</sup> are excellent examples.

Thus, biologically active compounds from nature promise to remain an important part of modern medicine into the foreseeable future.

### STRATEGIES FOR DRUG DISCOVERY FROM NATURE

How have drugs from nature been discovered? Various strategies have proved useful and these may be classified<sup>3</sup> into following four categories:

- (a) Study of human (and animal) biochemistry and physiology recognition of hormones, bioregulators, co-factors, enzymes and other biomacromolecules;
- (b) Random search—plants, microbes, marine flora and fauna;
- (c) Chemotaxonomic considerations; and
- (d) Ethnotherapeutics—folklore, traditional systems of medicine.

It will be worthwhile to briefly elaborate on these, and assess their worthiness for future drug discoveries.

Study of *mammalian chemistry* is a powerful tool for the development of new therapeutic agents. Use of hormones, vitamins and enzymes in health management has been the direct outcome of these efforts, which represent high mark of scientific achievement for the second quarter of this century. Amongst the hormones which were introduced in clinical usage during 1950-60, may be mentioned insulin, thyroxine, sex hormones, adrenocortical hormones and oxytocin. Special attention may be drawn to the steroidal sex and adrenocortical hormones,<sup>10,12</sup> the study of which dominated drug development from natural sources during the period 1950-65, when several semisynthetic analogues with superior therapeutic activity were developed and launched.<sup>13</sup> Since then, prostaglandins<sup>14</sup> and polypeptide hormones<sup>15</sup> have occupied dominating attention of several groups, because of much commercial potential. Polypeptide hormones are special targets for production by recombinant DNA and hybridomas technologies.<sup>16-17</sup> This powerful combination is all set to alter fundamentally biomedical research methodologies.

*Random search* for bioactive compounds from higher plants, lichens, bryophyta, microbes, and marine flora and fauna, has received extensive attention. Random search has proved most fruitful in the

discovery of antibiotics from a variety of microorganisms.<sup>18,19</sup> Extensive researches have also been carried out on plant materials for anticancer agents and though valuable leads have been obtained,<sup>20,21</sup> no drug has as yet reached the market *via* this strategy.

*Chemotaxonomic considerations* come into play once we know about the occurrence of a bioactive material in a plant or microorganism. Discovery of new antibiotics was greatly aided by these considerations.<sup>18</sup>

The fourth strategy is based on the *ancient medicinal knowledge* of various tribes, cultures and civilizations. If one looks at how the traditional plant drugs (cf Table I) came to be utilized in modern medicine, one will find that invariably, the starting point has been some reference to the use of that plant material as a particular cure in some folklore. Table III illustrates this point convincingly. Even after this classical plant drug discovery era, most of the important plant drugs have been discovered through folklore or traditional systems of medicine. Reserpine, the discovery of which in 1952, provided a distinct fillip to plant drug research, had its origin in *Ayurveda*: Its source *Sarpagandha* (*Rauwolfia serpentina*) is a well-recognized drug for the treatment of hypertension, insomnia and insanity, in the Ayurvedic system of medicine.<sup>25,26</sup> The discovery of two dimeric indole alkaloids, vinblastine and vincristine valuable for the treatment of Hodgkin's disease, lymphosarcoma and leukemia in children, from the periwinkle plant (*Catharanthus roseus*), is again the result of folklore remedy investigation, though in a different way. Work on this plant was taken up for its alleged hypoglycemic activity as per Jamaican folklore!<sup>3,25</sup> Podophyllum resin has been used by American Indians for treatment of cancer<sup>27</sup>; its active compound, podophyllotoxin, is the basis for etoposide and teniposide, anticancer agents currently being used against testicular cancer, small-cell lung cancer and lymphomas.<sup>28</sup>

From the above analysis it is clear that ethnotherapeutics provide a powerful and more effective strategy for the discovery of clinically useful compound from plant kingdom.

**Table III**  
*Ethnotherapeutics and traditional modern drugs*

| No | Drug                              | Basis of investigation   |
|----|-----------------------------------|--|
| 1  | <i>Codeine, morphine etc</i>      | Opium, the latex from <i>Papaver somniferum</i> used by ancient Sumerians, Egyptians and Greeks for treatment of headaches, arthritis, for inducing sleep. <sup>22,23</sup>              |
| 2  | <i>Atropine, hyoscyamine, etc</i> | <i>Atropa belladonna</i> , <i>Hyoscyamus niger</i> (henbane) were important items of Babylonian folklore <sup>24</sup>   |
| 3  | <i>Ephedrine</i>                  | Crude Drug <i>Ma-huang</i> (astringent yellow) derived from <i>Ephedra sinica</i> had been used by Chinese for respiratory ailments since 2700 B.C. <sup>24</sup>                        |
| 4  | <i>Quinine, etc</i>               | <i>Cinchona</i> spp. were used by Peruvian Indians for the treatment of fevers. <sup>22,24</sup>   |
| 5  | <i>Emetine</i>                    | Brazilian Indians and several other South American tribes used roots and rhizomes of <i>Ipecacuanha</i> ( <i>Cephaelis</i> spp.) to induce vomiting and cure dysentery. <sup>22,24</sup> |
| 6  | <i>Caffeine, etc.</i>             | These are major constituents of tea, coffee and cocoa used for preparing stimulating beverages in different parts of the world since many centuries ago.                                 |
| 7  | <i>Colchicine</i>                 | Use of <i>Colchicum</i> in the treatment of gout has been known in Europe since 78.A.D. <sup>24</sup>  |
| 8  | <i>Digoxin, etc</i>               | <i>Digitalis</i> leaves were being used in heart therapy in Europe during the 18th century. <sup>22</sup>  |

### ZHONG YAO—THE CHINESE SYSTEM FOR MEDICINE

The above conclusion gets reinforced when one surveys the results of modern investigations on Chinese medicine. Chinese medicine (*Zhong yao* in Chinese, *Kanpo* in Japanese) has been practiced in China and the neighbouring countries since more than 2000 years ago, and was apparently first systematized in *Shang Han Tsu Ping Lun*, a 16-volume compendium, possibly compiled by Chang Chung-Ching in the second century A.D (142-210). This work records eighty crude drugs, while

another later (456-536) compilation, *Shin-Nung Pen T'sao Ching*, of Tao Hung-Ching, describes 365 crude drugs.<sup>29</sup> Though the Chinese system of medicine, like other traditional systems of medicine had, in the past, been viewed with skepticism by the practitioners of Western system of medicine, there appears to be a healthy interest now.<sup>30</sup> The *Pharmacopoeia of the People's Republic of China*, issued in 1978 describes 882 crude drugs of which 637 are of plant origin. *Japanese Pharmacopoeia X* released in 1981, contains 102 plant drugs, of which only 29 are crude drugs used in Western medicine.<sup>30</sup> The incorporation of a rather large number of plant drugs in Chinese and Japanese recent pharmacopoeia is the direct consequence of modern investigations on Chinese medicine.

Chinese crude drugs have been the subject of intense modern scrutiny during the past two decades or so, mostly at the hands of Chinese.<sup>32-34</sup> Outstanding results have been obtained, essentially confirming many claims of the ancients. Several of these results have received international attention and acclaim. Thus, the traditional Chinese antimalarial drug, *qinghao* (aerial parts of *Artemisia annua*) has furnished a potent antimalarial, artemisinin (*qinghaosu*) (1), active against *Plasmodium vivax* and *P. falciparum* when given orally or intramuscularly.<sup>35,36</sup> This has been commercialised in China.<sup>32</sup> A simple derivative of artemisinin, called artemether (2), has been found to be superior for treatment of tertian and cerebellar malaria and is undergoing clinical evaluations.<sup>34,37,38</sup> Ginseng, which comprises roots of *Panax ginseng* is highly valued in the Orient as a tranquilizer, as an extender of memory, and as a general tonic. Modern investigations, mostly conducted by the Japanese have led to the isolation of a large number of triterpene glycosides (ginsenosides), mainly based on dammarane skeleton (e.g. ginsenoside Rb<sub>2</sub>, 3).<sup>29,33,39,40</sup> Several of these ginsenosides have been shown to exhibit a variety of pharmacological activity, such as increased DNA, RNA, protein and lipid synthesis in bone marrow cells, and suppression of cyclic AMP levels.<sup>39,41</sup> World market for ginseng and its products has been estimated (1980) at 1700 tonnes valued at around US \$ 70 million.<sup>42</sup>

The announcement by the Chinese, in 1978, that gossypol (4), the yellow (rather toxic) pigment of the cotton plant (*Gossypium* spp.) is an effective male contraceptive, was received by the international medical community with much excitement. Since then, the results have been



treatment of occlusive blood vessel diseases, such as cerebral embolism; lignans (e.g. schizandrin C, 8) from fruits of *Schizandra chinensis* as a protective agent against liver damage.<sup>32,34</sup> Trichosanthin, a well-characterized plant protein from the roots of Cucurbitaceae plant *Trichosanthes kirilowii* is being used in China as an abortifacient.<sup>35</sup>

Though several other examples from Chinese literature can be cited, the main purpose of citing these results, namely highlighting the importance of ethnotherapeutics, should have been served. Let us now have a look at our traditional system of medicine, that is Ayurveda, and evaluate its potential as a source for modern drug development.

## AYURVEDA AND DRUG DEVELOPMENT FROM PLANT SOURCES

Before proceeding further it appears appropriate to first present a brief introduction to *Ayurveda*. This may sound redundant as I am speaking to an audience of Indians in India. However, the unfortunate fact remains that many of us have not cared much to learn about our own heritage, and often look upon it with mistrust and disdain—the direct consequence of our system of education started by the British (understandingly) and perpetuated, since 1947, by our own people!

*Ayurveda* is the ancient Indian system of health-care, both physical and mental. The word *Ayurveda* is derived from *Ayus(r)*, meaning life, and *Veda*, meaning knowledge. Thus *Ayurveda* literally means "Science of life". This covers the art of living. Health, in *Ayurveda*, has been defined as a well-balanced metabolism plus a happy state of being. Disease has been considered four-fold, as emanating from (i) body (ii) mind, (iii) external factors, and (iv) natural intrinsic causes. The definition of treatment, in *Ayurveda*, covers a salubrious use of drugs, diets and practices.<sup>45</sup>

Pharmaceutics occupies an important place in *Ayurveda*. Medicinal preparations are invariably complex mixtures, being derived from plant and animal products, as also minerals and metals. Plants form dominant part of Ayurvedic pharmacopoeia. Earliest references to such plants are to be found in the *Rigveda* and *Atharvaveda*, dating back to second millennium B.C. *Charaka Samhita* (–900 B.C.) is the first recorded treatise on *Ayurveda*; it consists of eight sections divided into 150 Chapters, and describes 341 plants and plant products for use in medicine. Therapeutics (*Kayachikitsa*) is the hallmark of this work.<sup>46,47</sup> The next landmark in *Ayurveda* is *Sushruta Samhita* (–600 B.C.) with special emphasis on

surgery; it has six sections covering 186 chapters, and described 395 medicinal plants, 57 drugs of animal origin, and 64 minerals and metals as drugs. Sushruta, the Father of Surgery, lived and practiced surgery in Varanasi, some 2500 years ago.<sup>48,49</sup> The next important authority in *Ayurveda*, after Charaka and Sushruta, was Vagabhatta of Sind, who practiced about 7th Century A.D. His works, *Astanga Hridaya*, is considered unrivalled for principles and practice of medicine. *Astanga Hridaya*, is described in six sections covering 120 Chapters, and contains 7444 verses; the entire book is in verse.<sup>46</sup> Charaka, Sushruta and Vagabhatta are the *Vrihat Trya* (Powerful Triad) of *Ayurveda*. *Madhava Nidana*, consisting of 69 chapters and 1,552 verses, is the next important landmark. This work excels in diagnosis. It was written in 12th century by Madhava, the famous Vidyaaaranya of Vijayanagar. Sarangdhara (14th century), the author of *Sarangdhara Samhita* systematised Ayurvedic materia medica; this work consist of three parts, 32 chapters and 2,500 verses. The last celebrated writer on Hindu medicine is Bhava Mishra of Magadha, and his treatise *Bhava Prakasha* written ~ 1550 is held in high esteem by the modern Ayurveda practitioners. It has three sections containing 10,831 verses; approximately, 470 medicinal plants are mentioned.<sup>46,50</sup> Madhava, Sarangdhara and Bhava Mishra have been referred to as the *Laghu Trya* (Junior Three) of *Ayurveda*. Besides these monumental treatises, a rather large number of (>70) *Nighantu Granthas*) or Pharmacy lexicons have been written, mostly between 7th and 16th century.<sup>51,52</sup> *Raja Nighantu* by Narhari Pandita, and *Madanpala Nighantu* by Madanpala are considered as masterpieces on medicinal herbs.<sup>46</sup>

Eight branches of medicinal knowledge (*Astanga*) have been recognized in *Ayurveda* from ancient times. The lists given by Charaka, Sushruta and Vagabhatta are identical, though different in order.<sup>49</sup> Table IV gives this classification which is relevant to this day. Drugs have been classified on the basis of their physiological action.<sup>53</sup> For example, Charaka has divided plant dugs into fifty groups according to the actions of their decoctions. Table V gives part-list of this classification,<sup>47</sup> by way of illustration.

This introduction should suffice to emphasize that Ayurveda was a well-organised science of its times and this store-house of knowledge and information offers a unique opportunity for further exploration by all the powerful and subtle techniques of modern science. In the present context the Ayurvedic plant drugs deserve to be explored for modern drug development.



**Table IV**  
*Eight branches of Ayurveda*

| No. | Sanskrit designation    | Branch  | Remarks |
|-----|-------------------------|---|---------|
| 1.  | <i>Kayachikitsa</i>     | Therapeutics: all aspects   |         |
| 2   | <i>Shalakyatantra</i>   | Diseases of eyes, ears, nose, tongue, oral cavity, throat   |         |
| 3   | <i>Shalyatantra</i>     | Surgery   |         |
| 4   | <i>Kaumarabhritya</i>   | Paediatrics: anti-natal and post-natal baby care, and care of mother before conception and after pregnancy: diseases specific to children and their treatment |         |
| 5   | <i>Agadtantra</i>       | Toxicology . deals with poisons of various types and their antidotes; environmental and water pollution   |         |
| 6   | <i>Bhutvidya</i>        | Psychiatry  |         |
| 7   | <i>Rasayana</i>         | Knowledge of tonics and processes for arresting the process of mental and physical decay, rejuvenation  |         |
| 8   | <i>Vasikaranatantra</i> | Knowledge of virilifics, dealing with lost or diminished virility, potency and procreative ability  |         |

**Table V**  
*Classification of plant drugs (Charaka)\**

| No**. | Sanskrit name           | Group                 | Remarks |
|-------|-------------------------|-----------------------|---------|
| 1     | <i>Jivyaniya</i>        | Promoting longevity   |         |
| 3     | <i>Lekhaniya</i>        | Anti-obesity          |         |
| 6     | <i>Dipaniya</i>         | Promoter of digestion |         |
| 7     | <i>Balya</i>            | Promoting strength    |         |
| 8     | <i>Varnya</i>           | Complexion-promoting  |         |
| 15    | <i>Krimighna</i>        | Anthelmintic          |         |
| 17    | <i>Stanyajanana</i>     | Galactagogue          |         |
| 23    | <i>Vamanopaga</i>       | Emetic                |         |
| 24    | <i>Virechanopaga</i>    | Purgative             |         |
| 35    | <i>Mutravirechaniya</i> | Diuretic              |         |
| 36    | <i>Kasahara</i>         | Antitussive           |         |
| 38    | <i>Svayathuhara</i>     | Antiinflammatory      |         |
| 39    | <i>Jvarahara</i>        | Febrifuge             |         |
| 47    | <i>Vedanasthapana</i>   | Analgesic             |         |
| 50    | <i>Vayahsthapana</i>    | Anti-ageing           |         |

\* Listing is only partial

\*\* Numbering as per Charaka

It may be mentioned that the Ayurvedic texts were translated into Greek (300 B.C), Tibetan and Chinese (300 A.D), Persian and Arabic (700 A.D), and languages of other Asian people.<sup>54-56</sup>

### MODERN INVESTIGATIONS ON AYURVEDIC DRUGS

With the introduction of western scientific methods in India, Ayurvedic drugs and other Indian plants with alleged medicinal properties came under some sort of scrutiny. Many plants and their products came to be examined by chemists and medical people, though without any significant collaboration between the various disciplines. A fairly comprehensive account of these researches covering approximately 100 plant drugs, is given in Col. Ram Nath Chopra's now classic, *Indigenous Drugs of India* (1958).<sup>57</sup> These investigations did result in the confirmation of some attributes of these drugs.

Around 1964 systematic screening of Indian flora was undertaken at CDRI, Lucknow,<sup>58</sup> and CIBA-Geigy Research Centre, Bombay.<sup>59</sup> One really does not know the outcome of CIBA programme, as for obvious reasons, the results have not been made public. CDRI, on the other hand, regularly published the results of its findings. During a period of some 20 years, CDRI screened approximately 2500 plants for a variety of activities. Important leads appear to have been obtained and the results have been summarized in a recent publication.<sup>60</sup>

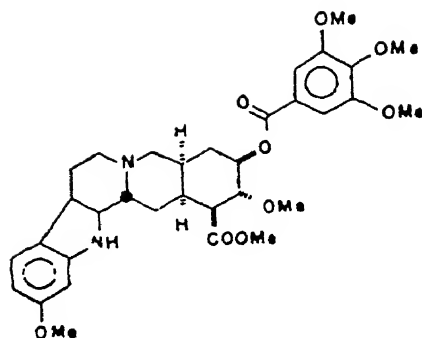
In 1961, the then Central Council of Ayurvedic Research (CCAR), Government of India, arranged a conference of reputed Ayurvedic Vaidyas to prepare a list of effective Ayurvedic plants. A list of 190 drugs, thus, emerged. Next, a Composite Drug Research Scheme (CDRS) was formulated by the Ministry of Health in collaboration with the Indian Council of Medical Research (ICMR), the then CCAR, and the Council of Scientific and Industrial Research (CSIR). This was launched in 1964, and was technically under the administrative control of ICMR till April 1970, when the newly created Central Council for Research in Indian Medicine & Homeopathy took charge of the scheme. Under this scheme nine circuits were set up in different parts of the country. Each circuit comprised four units, namely clinical (both Ayurvedic and Modern), pharmacognosy, chemical and pharmacological.<sup>61</sup> This scheme appears to have operated through the seventies. Results of these investigations have been published in two books.<sup>62,63</sup> besides in articles in other professional journals.

Besides the above major efforts, work on Indian medicinal plants has also been carried out in various universities and research institutes, and results reported in professional journals. The two above-mentioned books<sup>62,63</sup> also summarise these findings.

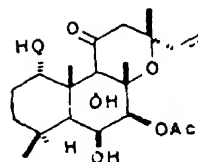
How does one evaluate the impact of these efforts spanning several decades? To my mind, the study of medicinal plants has clearly one objective, namely, development of improved and better drugs for health-care. Training of personnel, discovery of new chemical entities or recognition of new concepts, though important in their own right, can only be considered as spin-off and incidental to the main purpose. Viewed thus, the impact of these efforts has been rather dismal. Nothing from the above major programmes has as yet reached the market!

However, it may be pointed out that in 1953, reserpine (9) briefly alluded to earlier, and a minor alkaloid of Ayurvedic drug *Sarpagandha*, was introduced as an antihypertensive drug in modern medicine.<sup>64</sup> Though, the starting point for this development was a publication<sup>65</sup> by Dr R J Vakil of Bombay, the credit for successful development of a modern drug from this lead goes entirely to CIBA, Switzerland. A second contribution from *Ayurveda* is the very recent marketing of a hypolipaeamic drug based on *guggulu*, the gum-resin of *Commiphora mukul*.<sup>66</sup> I shall shortly discuss this in some detail. Besides this, I must also mention forskolin (colenol, 10), which was isolated independently by two different groups, one working at CDRI<sup>67,68</sup> and the other at Hoechst Research Centre, Bombay.<sup>69</sup> This compound activates hormone sensitive adenylate cyclase, thus potentiating cyclic AMP levels and providing a new mechanism to potentiate hormonal responses. Forskolin is a potent hypotensive agent, and also appears to be a valuable agent for the treatment of glaucoma. Several other therapeutic applications have been visualized.<sup>70-72</sup> Forskolin was isolated from roots of *Coleus forskohlii* (Syn. *C. barbatus*), a plant not mentioned in classical Ayurvedic literature. As a matter of fact roots of this plant (Gujarati, *Garmal*) are pickled and consumed in Gujarat.<sup>73</sup>

It may be instructive at this stage to analyse the reasons for this rather dismal performance. Development of a marketable drug is a complex business requiring an above-critical outlay of inputs (personnel, materials, expertise) and an effective organizational set-up. Judged from this angle, work in the universities, medical colleges and small institutes can never lead to fruition. No wonder India missed discovery of reserpine.



Reserpine

9

Forskolin

10

The CDRS, though well-conceived, proved impractical, as there was, despite best efforts, little co-ordination for several reasons mentioned elsewhere.<sup>74</sup> Though many Ayurvedic claims could be confirmed, the leads were not pursued further. Work at CDRI should have led to more fruitful consequences. Though, important leads were noted, somehow they could not be developed further, possibly because work undertaken was of a vast magnitude, without proper focus.

How does Ayurvedic materia medica, to which we have made some detailed references earlier emerge from all this? Does starting with a plant, reported in *Ayurveda* for a particular application, offer any additional or special advantage for drug development? A clear cut answer is not forthcoming from the above. While working as part of "Circuit-I" of CDRS, at NCL, Poona, we had occasion to investigate *shatavari* (*Asparagus racemosus*) and *punarnava* (*Boerhaavia diffusa*), and were happy to see confirmation of ancient claims. This reinforced my earlier views that Ayurvedic drugs which have withstood the test of time must have something in them. The (late) Dr C Dwarkanath, who was overseeing the work of CDRS, suggested that we also work on *guggulu*, which according to him, is valued in *Ayurveda* as an anti-inflammatory drug and for correcting lipid disorders. He was all the more keen, since according to general screening at CDRI, nothing worthwhile had been found in this drug, but he felt convinced that a different approach than that adopted at CDRI was called for.<sup>75,76</sup> He felt strongly, and I fully concurred with him, that while screening Ayurvedic drugs we should specially look for the

activity for which it is recommended in Ayurveda. Soon after that, I met Dr Nitya Anand (CDRI) and he readily agreed to collaborate in this work. This offered a good opportunity to answer the question which I posed earlier: does any lead from Ayurveda offer an advantage in drug development? From the story I am about to unfold an unequivocal yes did emerge!

### GUGGULU, AN ANCIENT AYURVEDIC DRUG

*Guggulu* is the gum-resin exudate of a small tree, *Commiphora mukul* (Syn. *Balasamodendron mukul*) belonging to the family Burseraceae. This tree is endemic to India and is found wild in semi-arid regions of Rajasthan, Madhya Pradesh, Gujarat and Karnataka. In its natural setting the tree remains essentially denuded of its foliage for most of the year. On injury, the plant exudes a yellowish gum-resin, which soon solidifies to an agglomerate of tears or stalactitic pieces with balsamic odour. These trees are tapped commercially, the average yield of gum-resin per tree being 700–900g/year, and the collection is done during the cold season. This is the *guggulu* of commerce.<sup>77</sup>

*Guggulu* is highly valued in *Ayurveda*. It is mentioned in *Atharva Veda*. It is specially recommended for the treatment of, lipid disorders and rheumatoid arthritis (Sushruta, Vagabhatta). Fig. 1 reproduces the Sanskrit *Shloka*, given in *Bhava Prakasha*, describing the various attributes of *guggulu*. Thirtythree compound preparations containing *guggulu* as a component are described in *Bhava Prakasha*.<sup>78</sup>

Modern pharmacological studies on the crude drug and some of its fractions had tended to support the claims of Ayurveda. Thus, anti-arthritis and antiinflammatory activities were confirmed by Gujaral and co-workers.<sup>79,80</sup> Dwarakanath, Satyavati and co-workers were able to establish the value of *guggulu* in lipid disorders, especially as a hypocholesterolemic and hypolipemic agent.<sup>81–84</sup> These findings were soon confirmed by other workers.<sup>85,86</sup> Thus, when we were introduced (1969) to *guggulu*, sufficient basis existed for undertaking detailed chemical and biological investigations.

Chemical work was soon undertaken at NCL, Poona and was later continued at Malti-Chem Research Centre, Vadodara. All biological work was carried out at CDRI, Lucknow under the leadership of Dr Nitya Anand and Dr Swarn Nityanand. Important results were periodically reported in seminars<sup>88,89</sup> or other publications.<sup>90–94</sup> During the course of

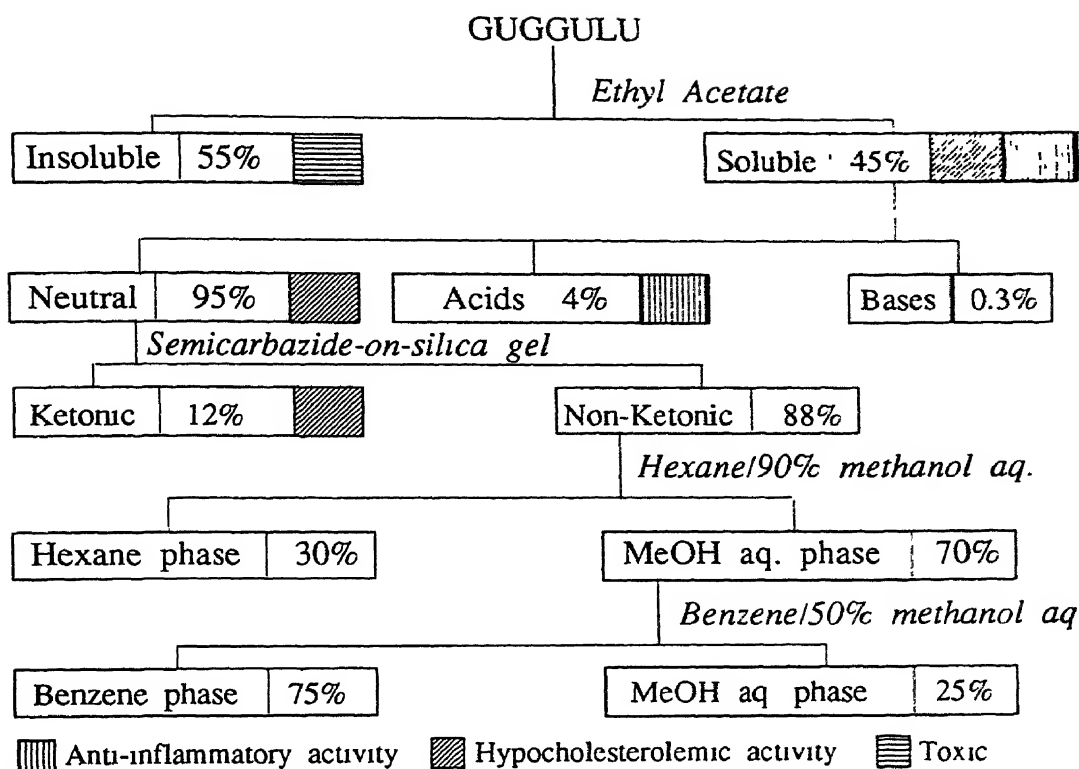
these investigations, a significant number of publications on biological activities of *guggulu* or its fractions were published by other groups.<sup>62,95</sup>

गुग्गुलुर्विशदस्तिक्तो वीर्योष्णः पित्तलः सरः ।  
 कषायः कटुकः पाके कटुरुक्षो लघुः परम् ॥  
 भस्मसंधानकृद्वृष्यः सूक्ष्मः स्वर्यो रसायनः ।  
 दीपनः पिच्छिलो बल्यः कफवातघ्नपाणचो ॥  
 मेहो मेहांश्च वातास्त्रक्लेदकुष्ठाममारुतान् ।  
 पिडिकान्ग्रन्थिशोफार्शोगण्डमालाकृमीञ्जवेत् ॥  
 माधुर्याच्छमयेद्भातं कषायत्वाच्च पित्तहा ।  
 तिक्तत्वात्कफजित्तेन गुग्गुलुः सर्वदोषहा ॥

FIG 1 Shlokas from *Bhava Prakasha* giving attributes of *Guggulu*.

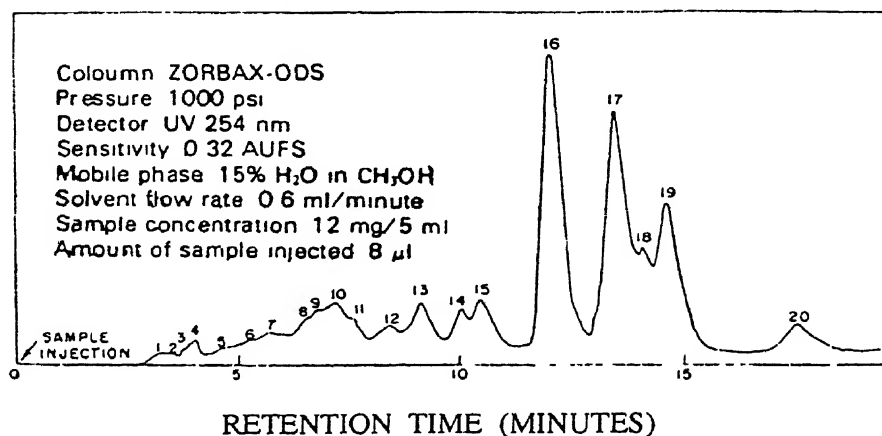
#### *Segregation of Guggulu and Isolation of Compounds with Hypolipemic Activity*

After considerable effort, involving biological screening at each stage, a viable separation scheme was evolved and this is shown in Fig. 2. The ethyl acetate soluble portion, which constitutes ~ 45 per cent of the gum-resin carries both the hypocholesterolemic and anti-inflammatory activities. The insoluble part, which represents the carbohydrate gum portion of the gum-resin studied earlier by Bose and Gupta,<sup>96</sup> was found to be toxic and hence, was not investigated further. The ethyl acetate-soluble part was chemically exceedingly complex. It was, next, separated into bases, acids and a neutral cut. The neutral fraction carries practically all hypocholesterolemic activity, while the acidic cut showed significant anti-inflammatory activity. It was soon found that the neutral portion contained several ketones, which had significant lipid-lowering activity. A simple method using semicarbazide-on-silica gel was developed<sup>97</sup> to effectively separate these ketones, which constituted ~ 12 per cent of the total neutral part. The non-ketonic portion was essentially devoid of any lipid-lowering activity.

FIG 2 Segregation of *Guggulu*

Next phase of the research was aimed at separating the ketonic mixture into individual compounds and establishing which of these are responsible for the lipid-lowering activity. As judged from thin-layer-chromatography (TLC) and high-performance-liquid-chromatography (HPLC; Fig. 3), the ketonic fraction was quite complex. HPLC shows a minimum of twenty components. Actually, the number can be much more. That this indeed was so, became clear when we later learnt that HPLC peak 16 consisted of four compounds. By systematic column chromatography aided by TLC, we finally succeeded in isolating several pure compounds,<sup>92a,94</sup> including those responsible for lipid lowering activity. We consider this a fortuitous achievement, as isolation of useful biologically active compounds from natural materials is often than not beset with many hurdles and imponderables. Structures of these compounds were established with the aid of modern spectral techniques coupled with degradative or synthetic work as found necessary. These structures are shown in Fig. 4. The approximate concentration, as ascertained by actual isolation of each of these ketones in the total gum-resin, is also given. From Fig. 3, it is clear that these compounds constitute 60 per cent of the total ketone fraction. Compounds 11, 12, 13 have been known earlier. Compound 11, now named *Z*-guggulsterol, had been

known from synthesis, but was encountered in nature for the first time in *guggulu*. It is interesting to know that the other two minor constituents (12,13) of the gum-resin have been reported earlier as components of defence secretions of certain water beetles. Thus, compound 12 is present in the defence secretion of *Acilius sulcatus*, while its 20-epimer (13) is elaborated by its cousin *Hybius fenestratus*.<sup>98</sup>



|      |   |    |                                  |
|------|---|----|----------------------------------|
| 1-15 | Unidentified  | 18 | unidentified                     |
| 16   | : Z-guggulsterol, (20S - 20-hydroxy-4-pregnen-3-one, guggulsterol-VI, | 19 | (20R)-20-hydroxy-4-pregnen-3-one |
| 17   | Z-guggulsterone   | 20 | guggulsterol-1 E-guggulsterone   |

FIG 3 High-performance-liquid-chromatography of ketone fraction of *Guggulu*

Compounds 15, 16 which constitute some 2 per cent of the gum-resin are the major components of the ketonic fraction, and are responsible for the lipid-lowering activity of *guggulu*. These compounds, though isolated from nature for the first time, had been known synthetically much earlier.<sup>99</sup> As is the usual practice we have given these compounds the trivial names, Z-guggulsterone (15) and E-guggulsterone (16).

The remaining three ketones (14, 17, 18) are new compounds and their structures were confirmed by partial synthesis. Final stereochemical details for guggulsterol-I (18) were established by X-ray crystallography.<sup>100</sup>

The lipid-lowering activity of Z- and E-guggulsterones has been mentioned earlier. However, since the total ethyl acetate extract of gum resin has comparable activity, it follows that other constituents present in the ethyl acetate extract may be exerting a synergistic action on the



biological activity of these compounds. In view of this a detailed chemical analysis of other fractions (Fig. 2) was undertaken.

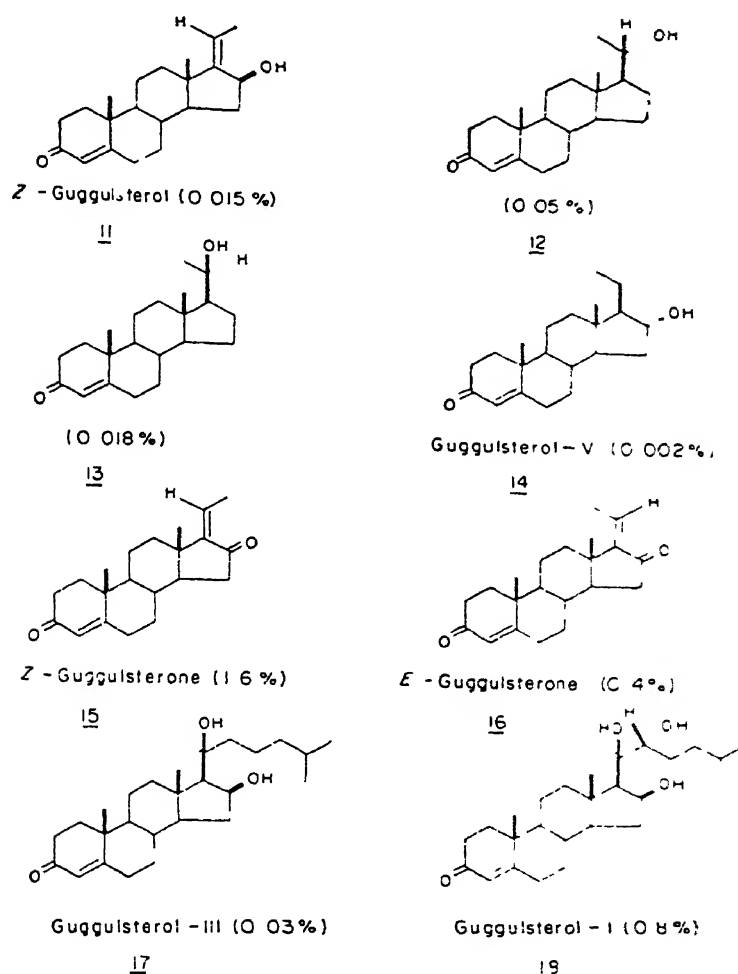
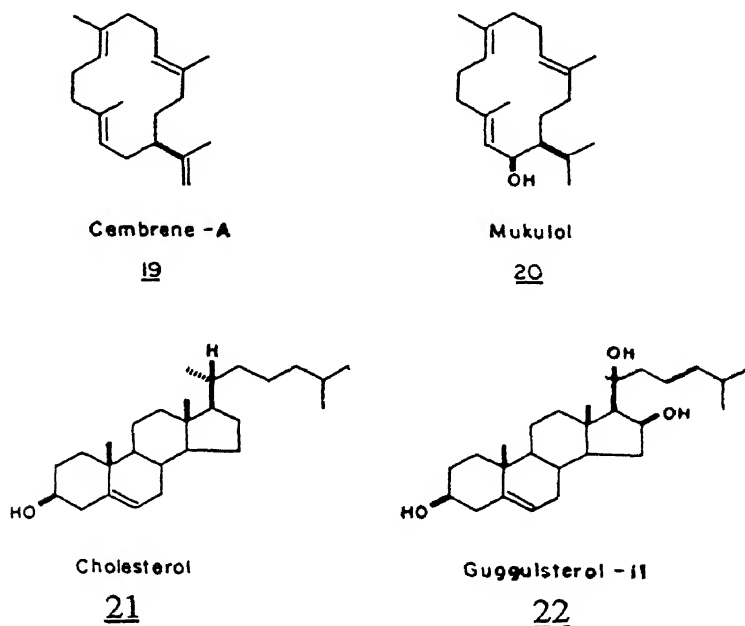


FIG 4 Ketones isolated from *Guggulu*

### Other Constituents

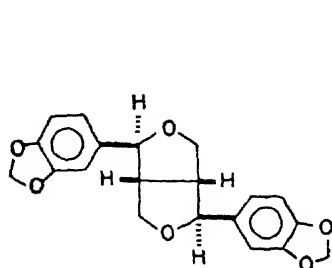
The hexane phase material constitutes ~9 per cent of the gum-resin, and consists essentially of diterpenoids, besides a small percentage of cholesterol (21) and other unidentified compounds. The diterpenoids, consist of a new diterpene hydrocarbon, cembrene-A (19 ; ~10 per cent), and its related alcohol, mukulol (20; ~40 per cent), besides other related minor terpenes, which appear to be artifacts.<sup>92b,93,101</sup> Cembrene-A is the

simplest cembrenoid arising from 1,14-cyclization of geranyl-geranyl pyrophosphate, the biogenetic precursor of diterpenoids. Structure of this compound was arrived at by spectroscopic methods and confirmed by a chemical correlation with (+)-limonene. It is interesting to note that the trail-marker of Australian termites, *Nasutitermes exitiosus* has been shown to have the same structure.<sup>102</sup> When a modified procedure for separation was employed, another steroid, designated guggulsterol-II (22) could be isolated. Its structure was arrived at by spectral methods and confirmed by partial synthesis.<sup>92a</sup>

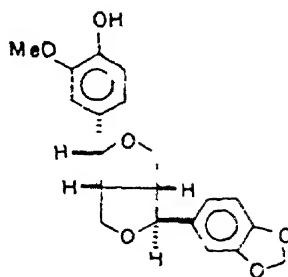


Material contained in the benzene phase (Fig. 2) constitutes ~ 14 per cent of gum-resin. Four lignans (23-26), of which two (23,24) were known, have been isolated from this cut. The two new lignans, named guggullignan-I (25) and guggullignan-II (26) were characterized by spectral methods.<sup>103,104</sup> A series of long-chain polyols esters were also isolated. These esters are derived from homologous polyols and ferulic acid. The polyols were recognized as tetrols of general structure 27 in which C<sub>20</sub> and C<sub>18</sub> compounds predominated. Configuration of these polyols has been established by synthesis starting with D-glyceraldehyde.<sup>104,105</sup> These polyols which have been given the generic name 'guggultetrol', have all been found to have the D-xylo configuration. Structure 28 is a typical representative of these compounds. This is the first report on the occurrence of such compounds in nature, and they

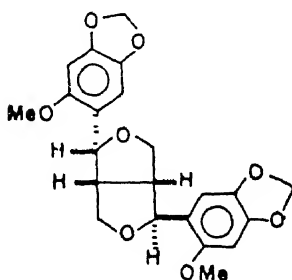
appear to be important because of their structural similarity with the biologically important phytosphingosines (29). Phytosphingosines are present in cerebrosides and gangliosides.<sup>106</sup>



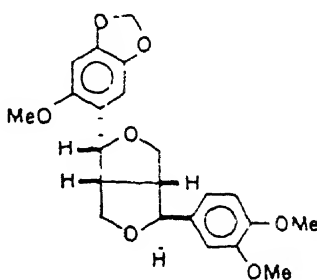
Sesamin

23

Pluviatilol

24

Guggullignan - I

25

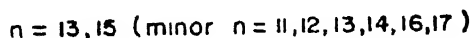
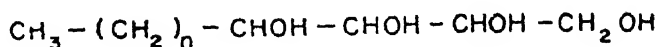
Guggullignan - II

26

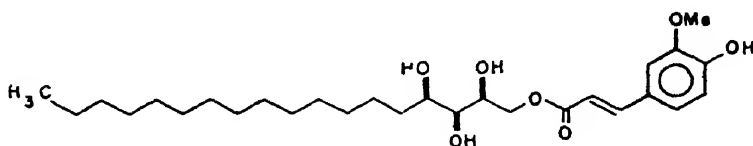
### *Gugulipid, a Modern Ayurvedic Drug*

Quite early in the course of this work it was concluded, that instead of using pure Z- and E-guggulsterones for developing a hypocholesterolemic/hypolipemic drug, it will be better if a standardized ethyl acetate extract of *guggulu* is developed for clinical use. This decision was motivated, firstly, on purely commercial considerations so that the product will be economically viable and competitive, and secondly, because of the fact that other components of this ethyl acetate extract appear to have significant synergistic effect. Thus, detailed pharmacological and toxicological studies and various phases of clinical trials have been conducted by CDRI on a standardized ethyl acetate extract (Fig. 2) of *guggulu*. This extract has been named 'gugulipid', and

has been standardized to contain 4.0g of Z-and E-guggulsterones per 100g of the extract. Estimation of these standards is readily done by HPLC.

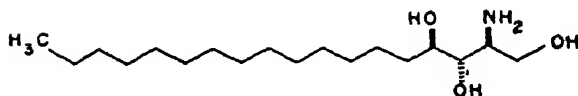


27



D - Xylo - Guggultetrol - 18 faralate

28



Phytosphingosine (C<sub>18</sub>)

29

Detailed pharmacological studies have been conducted on rats, rabbits and monkeys. Thus, in normal rats oral administration of gugulipid at 100mg/kg body weight daily for 30 days led to serum cholesterol lowering by 34 per cent and of serum triglycerides by 26 per cent. The corresponding values for 100mg/kg of clofibrate, a standard hypocholesterolemic drug, were 39 per cent and 30 per cent respectively. Fig 5 depicts the effect of gugulipid in hyperlipidemic rats (high fat diet), while Fig. 6 gives the results with hyperlipidemic monkeys; the treated monkeys showed 50 per cent and 30 per cent decrease in low-density and very low density lipoproteins respectively. In a similar study with hyperlipidemic rabbits (high-fat diet) oral ingestion of gugulipid at 50 mg/kg body weight daily for 8 weeks caused a fall in serum cholesterol by 40 per cent and triglycerides by 30 per cent.<sup>107</sup>

Gugulipid displays mild anti-inflammatory activity, no C.N.S. or diuretic activity. In mice the LD<sub>50</sub> value was 1600mg/kg both oral or

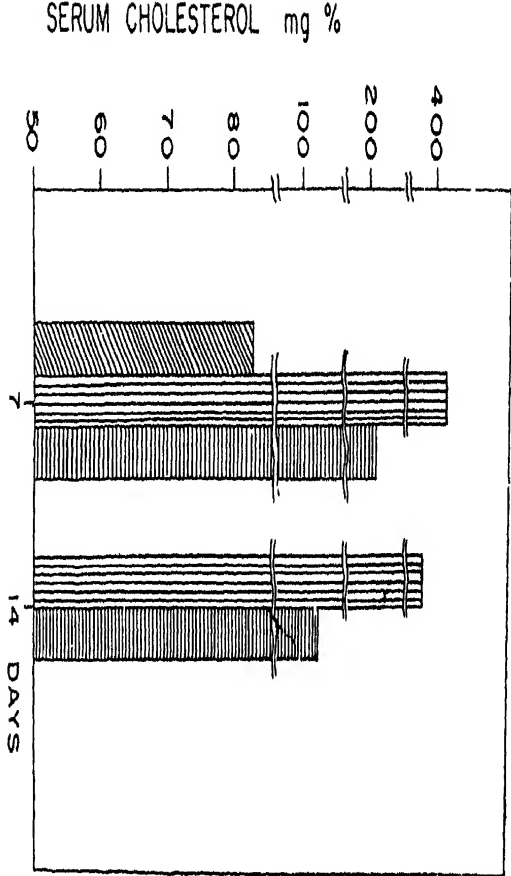


FIG 5 Guggulipid in hyperlipidemic rats

--- CONTROL  
--- 60 mg./kg  
--- 120 mg./kg.

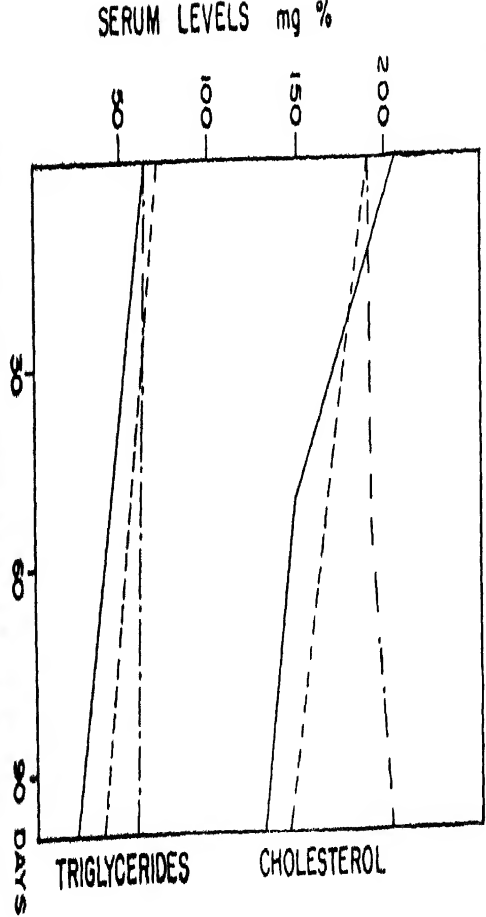


FIG 6 Guggulipid in hypolipidemic monkeys

intraparenteral. Subacute and chronic toxicity studies were carried out in rats, dogs and monkeys; no adverse effects were noticed. Gugulipid was also shown to be devoid of any teratogenic or mutagenic effects.

Phase I, II and III clinical studies have been carried out on human volunteers. In Phase III, for example, a total of 87 hyperlipidemic patients, male and female, in the age-group 38-48 years participated. The studies were double blind and were conducted in Bangalore, Bombay, Jaipur and Lucknow. The patients received 1500 mg/day of gugulipid or clofibrate in three divided doses for 12 weeks. The results showed that gugulipid is a safe and effective lipid-lowering agent comparable in efficacy to clofibrate, but with better compliance. Thus, for example, in Bangalore trials, the mean reduction was 22 per cent and 14.5 respectively for cholesterol and triglycerides in case of gugulipid, while the corresponding figures for clofibrate were 18 per cent (cholesterol) 14 per cent (triglycerides). The average initial levels of cholesterol and triglycerides in the patients were 220 mg per cent and 150mg per cent respectively.<sup>107</sup>

Gugulipid was cleared by Drug Controller (India) in June 1986 for marketing. It is now (Nov. 1987) being manufactured and marketed by CIPLA, Bombay.

### SHATAVARI

I would also like to discuss briefly the results we have obtained with another Ayurvedic crude drug, *shatavari*. *Shatavari* of commerce comprises dried, decorticated roots of *Asparagus racemosus*.<sup>62,108</sup> Several therapeutic attributes have been claimed in the classical Ayurvedic literature. Fig. 7 shows two such quotations pertaining to its use in cases of threatened abortion and as a galactagogue.<sup>109,110</sup> Preliminary pharmacological screening of various extracts of the crude drug appeared to confirm these claims.<sup>111-113</sup>

After considerable experimentation involving biological screening of various extracts for antioxytocin activity, scheme finally evolved for isolation of active fractions is shown in Fig. 8. At this stage, it should be mentioned that Professor B B Gaitonde, who collaborated in this project considered specific antioxytocin activity (rat, guineapig, rabbit uteri) as a good indicator of anti-abortifacient activity. This separation scheme led to the isolation of a) (glycosidic fraction having specific antioxytocin activity. This material was found to be a complex mixture of at least nine compounds. Of these, only one glycoside named shatavarin-I, showed

specific antioxytoec activity, both *in vitro* and *in vivo*; the compound showed no progestational or esterogenic activity.

शतावरी गुरुः शीता तिक्ता स्वाद्वी रसायनी ।  
मेधाग्नि स्निग्धा नेत्र्या गुल्मातिसारजित् ॥  
शुक्रस्तन्यकरी बल्या वातपित्तास्त्रशोथजित् ।  
महाशतावरी मेध्या हृद्या वृष्या रसायनी ॥  
(भास्कराकाश)

शतावरी सिद्ध घृत (पुंसवनं/गर्भदं)

(अ. हृदय उत्तरस्थान)

FIG 7 Sanskrit *Shlokas* giving medicinal attributes of *Shatavari*

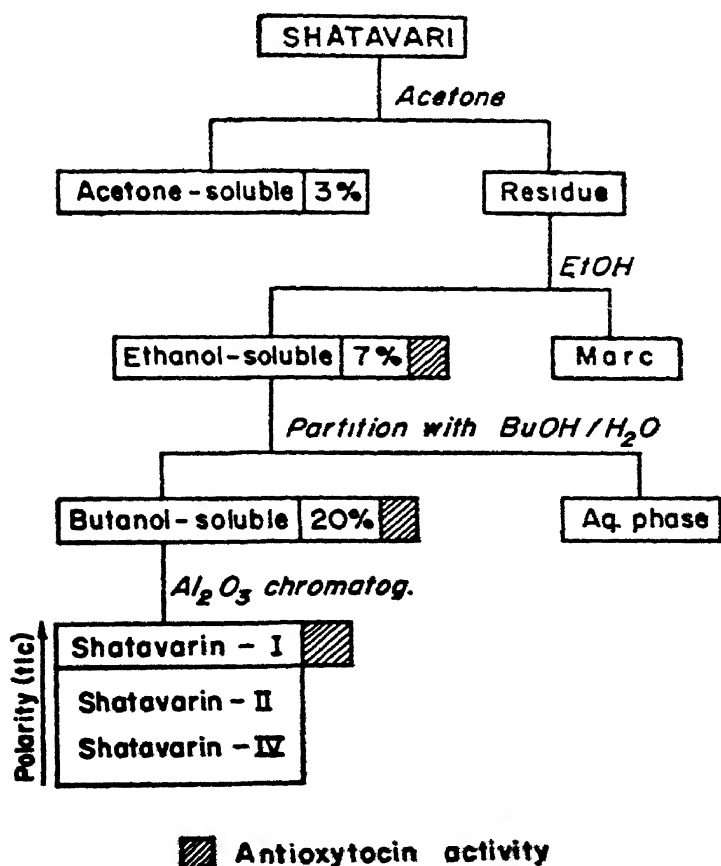
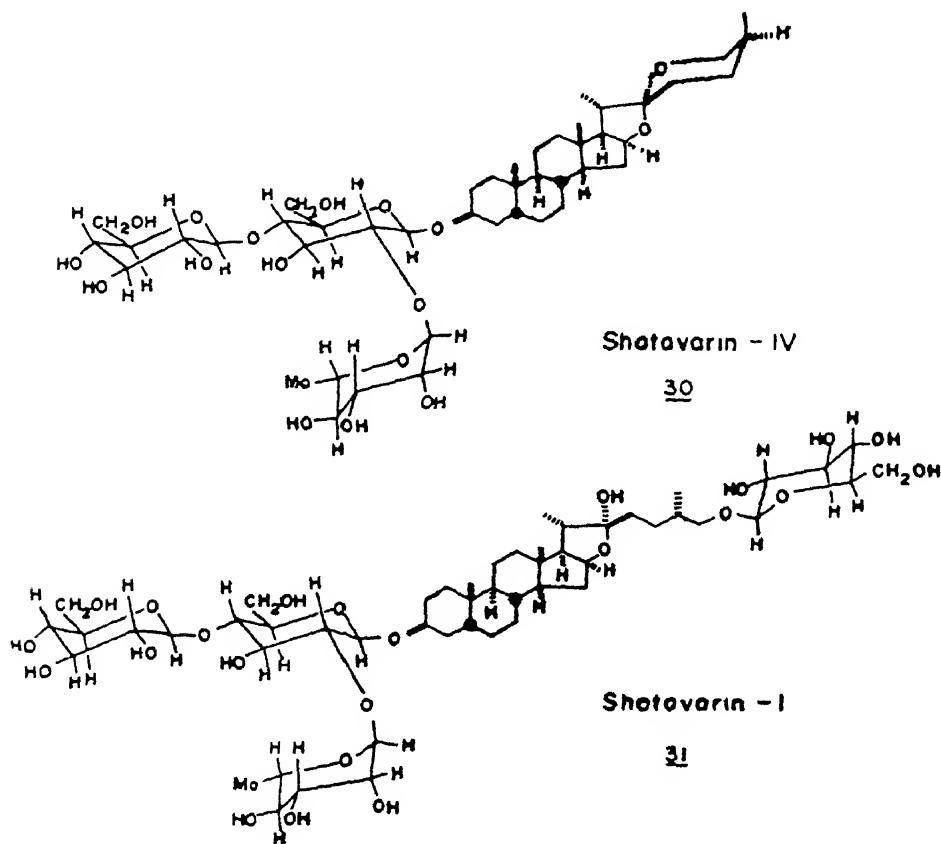


Fig 8 Scheme for the isolation of an active principle from *Shatavari*.

Structure of Shatavarin-I (31) has now been worked out essentially on the basis of its spectral characteristics and enzymic hydrolysis to shatavarin-IV(30).<sup>114,115</sup>

Thus, this study again turned out to be confirming ancient claims of Ayurveda.



## HARNESSING AYURVEDA FOR DRUG DEVELOPMENT

The material presented thus far puts Ayurvedic *materia medica*, as judged by modern parameters, in a reasonably acceptable position. Over the past two decades or so, a lot of data on the pharmacological, biological and therapeutical activity of many Ayurvedic single drugs, as well as compound preparations, have been published.<sup>116</sup> In any number of cases biological attributes of the ayurvedic drugs appear to have been demonstrated. By way of illustration, some examples are given in Table VI. Even though, much of these data might not have been collected under



ideal experimental conditions, a positive trend is more than evident. Considering all these, along with many as yet unexplored Ayurvedic drugs and folklore herbs, we have on hand a promising resource for development of useful drugs on modern lines. How does one go about achieving this?

**Table VI**  
*Some ayurvedic crude drugs with pharmacologically/  
therapeutically proven claims*

| No | Plant<br>(Botanical)              | (Sanskrit)          | Type of Activity                                 |
|----|-----------------------------------|---------------------|--|
| 1  | <i>Achyranthes aspera</i>         | <i>Apamarga</i>     | Diuretic <sup>62,117</sup>                       |
| 2  | <i>Acorus calamus</i>             | <i>Vacha</i>        | Tranquilizer <sup>62</sup>                       |
| 3  | <i>Adhodo vasica</i>              | <i>Vasa</i>         | Bronchodilator, Oxytocic <sup>118</sup>          |
| 4  | <i>Artemisia vulgaris</i>         | <i>Nagadaman</i>    | Cardiac tonic <sup>50,62</sup>                   |
| 5  | <i>Bacopa monnieri</i>            | <i>Brahmi</i>       | "Memory" <sup>62,119,120</sup>                   |
| 6  | <i>Boerhaavia diffusa</i>         | <i>Punarnava</i>    | Diuretic, Antiinflammatory <sup>62,119,121</sup> |
| 7  | <i>Butea frondosa</i>             | <i>Palasha</i>      | Anthelmintic <sup>62,119,122</sup>               |
| 8  | <i>Cassia fistula</i>             | <i>Aaragwadha</i>   | Cathartic <sup>62,110</sup>                      |
| 9  | <i>Cedrus deodara</i>             | <i>Devadaaru</i>    | Spasmolytic <sup>110,123</sup>                   |
| 10 | <i>Celastrus paniculatus</i>      | <i>Iyotishmati</i>  | "Memory" <sup>110</sup>                          |
| 11 | <i>Centella asiatica</i>          | <i>Mandukaparni</i> | "Intelligence" <sup>62,119</sup>                 |
| 12 | <i>Cissus quadrangularis</i>      | <i>Vajravalli</i>   | Healing of bone fractures <sup>62,119</sup>      |
| 13 | <i>Curcuma longa</i>              | <i>Haridra</i>      | Antiinflammatory <sup>110,119,125</sup>          |
| 14 | <i>Eugenia jambolana</i>          | <i>Jamboo</i>       | Hypoglycemic <sup>62</sup>                       |
| 15 | <i>Holarrhena antidysenterica</i> | <i>Kutaja</i>       | Antidysenteric <sup>63,110,126</sup>             |
| 16 | <i>Picrorhiza kurroa</i>          | <i>Katukaa</i>      | Antihepatotoxic <sup>63,127</sup>                |
| 17 | <i>Pueraria tuberosa</i>          | <i>Vidarikanda</i>  | Galactagogue <sup>110,128</sup>                  |
| 18 | <i>Sida rhombifolia</i>           | <i>Mahabala</i>     | Anabolic <sup>129</sup>                          |
| 19 | <i>Swertia chirata</i>            | <i>Kaurata</i>      | Febrifuge <sup>110,130,131</sup>                 |
| 20 | <i>Withania somnifera</i>         | <i>Ashwagandha</i>  | Anabolic <sup>110,130,131</sup>                  |

Before proceeding to answer the above question, one must tackle another one first, that is, if the Ayurvedic preparations are efficacious why not continue with these as such? To my mind, the answer is simple. Modern science and technology have provided us with unique tools, which must be used to further refine these preparations. No system, no knowledge can afford to remain static. Last word on any subject can never be written. With expanding knowledge, horizons of perception widen, generating new, valid points of view, which must be met. Ayurveda itself provides an excellent example of steady development till the medieval

period, when because of foreign onslaughts and political domination. Ayurveda became stagnant.<sup>46,134</sup> Thus, it is essential that Ayurvedic preparations get all the benefits of modern chemical, biological and pharmaceutical practices.

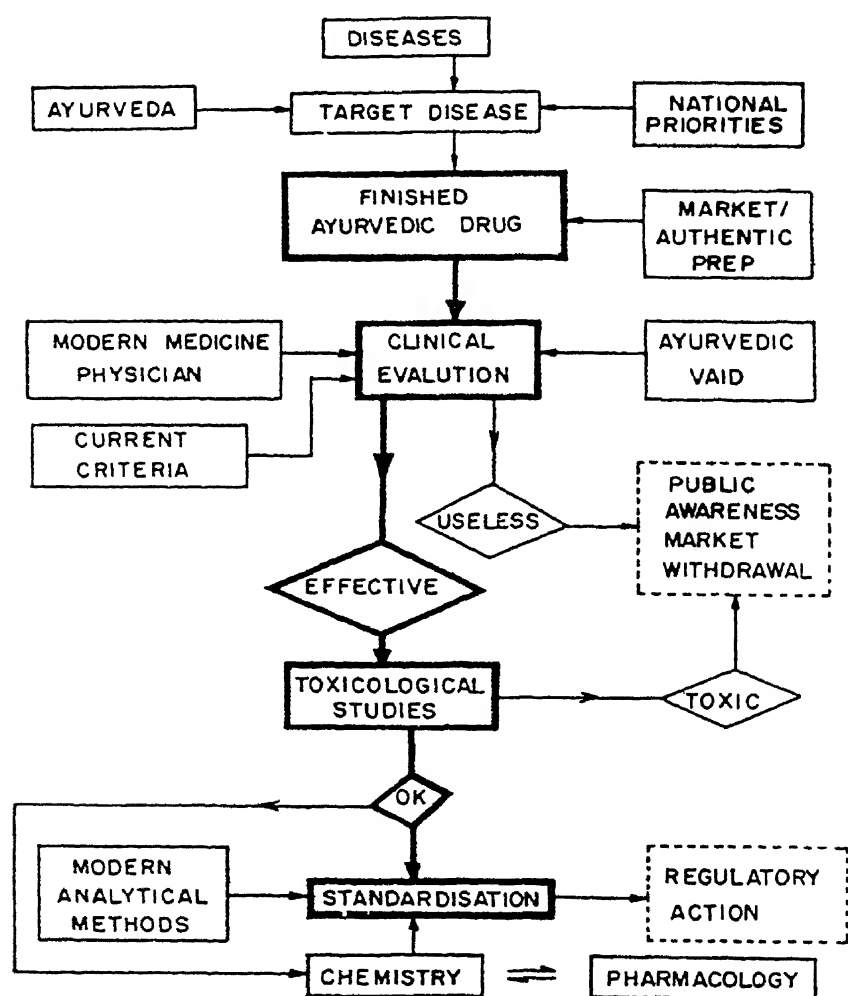


FIG 9 A plan for modern version of ayurvedic compound drugs.

Now, returning to the first question of harnessing Ayurveda for modern drug development, a different approach than that normally adopted for drug discovery from plant sources is called for. Since the Ayurvedic formulations, as well as its crude drugs have been in use in this country and the neighbouring countries for many many centuries, it would not be unethical to evaluate these materials clinically first. I have propagated this view, which formed the core of the Composite Drug

Research Scheme,<sup>76</sup> over the last many years.<sup>3,74</sup> This approach would also take care of the various imponderables, such as the question of right animal models, therapeutic action *via* prodrugs or by way of improving body functioning, or the question of synergism/antagonism. Such factors have often vitiated the outcome of drug discovery from natural sources. Efforts should be expended in two directions : (i) evaluation of therapeutic efficacy of standard Ayurvedic preparations and collecting sufficient scientific data to permit standardization, and (ii) search for new biologically potent entities from Ayurvedic single drugs to permit later, development of new drug as such or by providing a new model for development of a more effective synthetic analogue Figs. 9 and 10, which are self-explanatory, depict the various stages which will have to be gone through in this effort.

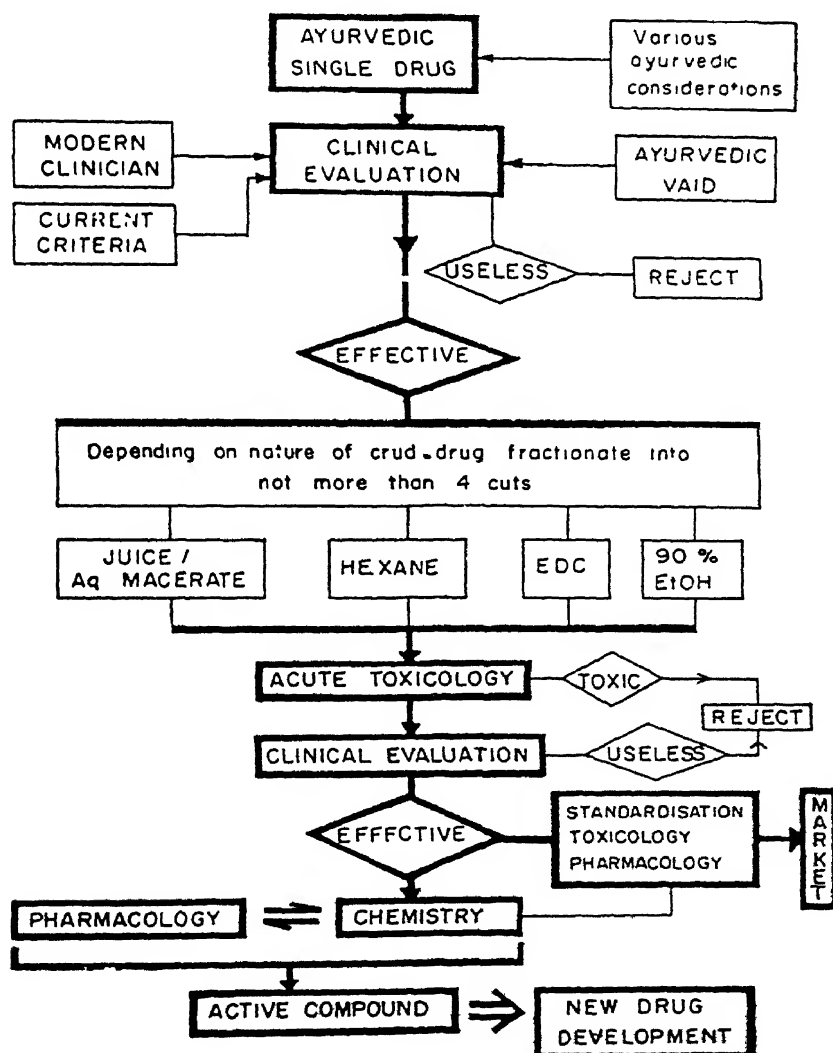


FIG 10 A plan for ayurveda-based modern drug discovery

In this suggested programme there are two crucial points which need elaboration. First is the selection of the target disease. In the final analysis, drug therapy is one. It is immaterial what the source of the drug is, so long as it is effective, safe and reasonably priced. Thus we should invest our time, effort and funds in those areas where Ayurvedic *materia medica* has something unique to offer. Hepatoprotective compounds, anti-ageing formulations, memory and intelligence enhancing products, compounds active against metabolic diseases, anabolics, antifertility compounds, rapid wound-healing products, immunopotentiators are some of the areas in which Ayurveda may have something worthwhile to offer. The second important point pertains to clinical evaluation of the drug. Both Ayurvedic and modern medicine physicians should be involved and should operate with full mutual understanding. The course of improvement in clinical manifestation of the condition should be judged using the latest criteria and standards. The objective would be to uncover something profound and not just a triviality.

In this endeavour the usual factors associated with drug supply from natural sources will have to be tackled along scientific lines.

The final success of the programmes delineated in Figs. 9 and 10 will have to be measured in economic terms. A therapeutic breakthrough will remain sterile unless it is rewarded by economic success.

### EPILOGUE

At the end, I wish to reemphasize the importance of Ayurveda in this part of the world and its potential in modern drug development. *Ayurveda* as we all know is still widely practiced in the Hindustan peninsula. The number of Ayurvedic practitioners, both registered and otherwise, in India is estimated<sup>134</sup> to be 2,32,227. It is imperative that benefits of modern science flow to *Ayurveda* to make it once more vibrant and, pulsating with new activity. Ayurvedic *materia medica*, I believe, represents a treasure-trove waiting to be explored by all the modern techniques. Even with the present state of development in medicine, satisfactory therapy is available only for about the third of all ailments.<sup>135</sup> May be *Ayurveda* has something to offer! Crores have been spent on the study of Indian medicinal plants and other formulations, but the outcome has been disappointing, chiefly because most of these studies were half-hearted, disjointed and ill-focussed. It is suggested that Government of India set up an effective modern facility dedicated to research in Ayurveda so as to

enable it to effectively intermesh with modern medicine. After all, in the ultimate analysis, there cannot be different versions of the same science.

### ACKNOWLEDGEMENT

I would like to take this opportunity to thank all my students and colleagues who worked on the chemistry of *guggulu* and *shatavari* and whose names appear in the references cited. Special thanks are due to Dr Nitya Anand and Dr Swarn Nityanand and their colleagues for the excellent collaboration which resulted in the emergence of an ancient Ayurvedic drug as a modern medicine.

### References

1. D S Kothari *Biogr Mem Fellows Natn Inst Sci India* 2 (1970) p 49
2. See e.g.: (a) R D McMurray In: *The Theory and Practice of Industrial Pharmacy* (Eds L Lachman H A Lieberman and J L Kanig) Varghese Publishing House Bombay (1985) p 736
3. Sukh Dev *Curr Sci* 52 (1983) 947
4. N R Farnsworth and A S Bingel In: *New Natural Products and Plant Drugs with Pharmacological, Biological or Therapeutical Activity* (Eds . H Wagner and P Wolff) Springer-Verlag Berlin (1977) p 1
5. L A Mitscher and A Alshmma *Ann Rep Med Chem* 15 (1980) 255
6. B M Blom in *Drug Discovery (Advances in Chemistry Series 108)* American Chemical Society Washington (1971) p 176
7. H Wagner and P Wolff (Eds) *New Natural Products and Plant Drugs with Pharmacological, Biological or Therapeutical Activity* Springer-Verlag Berlin (1977)
8. J L Beal and E Reinhard (Eds) *Natural Products as Medicinal Agents* Hippokrats Verlag, Stuttgart (1981)
9. N J de Souza B N Ganguli and J Reden *Ann Rep Med Chem* 17 (1982) 301
10. See e.g.: V Petrow in *Kirk-Othmer Encyclopedia of Chemical Technology* (Eds: M Grayson and D Eckroth) 12 Wiley-Interscience New York (1980) p 57
11. See. Eg. : W Gerhartz *et al* (Eds), *Ullmann's Encyclopaedia of Industrial Chemistry* A2 V C H Verlag Weinheim (1985) p 467
12. J A Hogg In: *Drug Discovery (Advances in Chemistry Series 108)* American Chemical Society Washington (1971) p 14
13. M J Green and B N Lutsky *Ann Rep Med Chem* 11 (1976) 149

14. N A Nelson R C Kelly and R A Johnson *Chem Engng News* Aug 16 (1982) 30
15. J Meienhofer *Ann Rep Med Chem* 11 (1976) 158
16. J J Monahan *Ann Rep Med Chem* 17 (1982) 229
17. R M Baum *Chem Engng News* July 20 (1987) 11
18. L H Conover in *Drug Discovery (Advances in Chemistry series 108)* American Chemical Society Washington (1971) p 33
19. M J Weinstein and G H Wagman (Eds) *Antibiotics* Elsevier Amsterdam (1978)
20. J M Cassady and J D Douros (Eds) *Anticancer Agents based on Natural product Models* Academic Press New York (1980)
21. A Aszalos (ed) *Antitumor Compounds of Natural Origin* C R C Press Boca Raton Florida (1982)
22. N Taylor *Plant Drugs that Changed the World* George Allen and Unwin London (1965)
23. W H Lewis and M P F Elvin-Lewis *Medical Botany* Wiley-Interscience New York (1977)
24. L W Codd *et al* (Eds) *Materials and Technology* V Longman London (1972) p 707
25. S M Kupchan In: *Drug Discovery (Advances in Chemistry Series 108)* American Chemical Society Washington (1971) p 1
26. B Mukerji *Medal Lectures II* Indian National Science Academy New Delhi (1984) p 973
27. J L Hartwell and A W Schrecker *Fortsch Chem Org Naturstoffe* 15 (1958) 83
28. W Gerhartz *et al* (Eds) *Ullmann's Encyclopedia of Industrial Chemistry A5*, VCH Verlag Weinheim (1986) p 21
29. S Shibata in. *Advances in Natural Products Chemistry* (Eds: S Natori N Ikekawa and M Suzuki) Kodansha Tokyo (1981) p 398
30. *Herbal Pharmacology in People's Republic of China* National Academy of Sciences Washington (1975)
31. H Hikino In : *Economic and Medicinal Plant Research* (Eds: H Wagner H Hikino and N R Farnsworth), 1 Academic Press New York p 53
32. Pei-Gen Xiao in *Natural Products as Medicinal Agents* (Eds: J L Beal and E Reinhard) Hippokrates Verlag Stuttgart (1980) p 351
33. H M Chang *et al* (Eds) *Advances in Chinese Medicinal Material Research* World Scientific Singapore (1985)
34. Zhu Da-yuan *Abstr Chinese Med* 1 (1987) 251

35. Xiao-Tian Liang In: *Advances in Chinese Medicinal Materials Research* World Scientific Singapore (1985) p 427
36. K Hillier *Drugs Fut* 6 (1981) 37
37. L J Bruce- Chwatt *Br Med J* 284 (1982) 767
38. *Ann Drug Data Rep* 5 (1983) 10
39. S Shibata O Tanaka J Shoji and H Saito in: *Economic and Medicinal Plant Research* (Eds H Wagner H Hikino and N R Farnsworth) 1 Academic Press New York p 218
40. O Tanaka R Kasi and T Morita *Absir Chinese Med* 1 (1986) 130
41. I Weinryb *Ann Rep Med Chem* 15 (1980) 190
42. *Markets for selected Medicinal Plants and their Derivatives* International Trade Centre UNCTAD/GATT, Geneva (1982) p 145
43. M R N Prasad E Diczfalusy In *Fertility and Sterility* (Eds R F Harrison and J Bonnar) MTP Press Lancaster U K (1983) p 255
44. D P Waller L J D Zaneveld and N R Farnsworth In *Economic and Medicinal Plant Research* (Eds H Wagner H Hikino and N R Farnsworth) 1 Academic Press New York p 87
45. S Sharma In: *Realms of Ayurveda* (Ed S Sharma) Arnold-Heinemann New Delhi (1979) p 11
46. P M Mehta In: *Realms of Ayurveda* (Ed S Sharma) Arnold-Heinemann New Delhi (1979) p 47
47. P Ray and H N Gupta *Caraka Samhita* National Institute of Sciences of India New Delhi (1965)
48. G D Singhal In: *Realms of Ayurveda* (Ed S Sharma), Arnold-Heinemann, New Delhi (1979) p 124
49. R C Majumdar In: *A Concise History of Science in India* (Eds: D M Bopse S N Sen and B V Subbarayappa) Indian National Science Academy, New Delhi (1971) p 213
50. G Pandey (Ed) *Bhava Prakash* Chaukhamba Vidyabhavan Varanasi (1960)
51. A S Paranjpe G S Pendse and V A Bedekar (Eds) : *Laghu-Nighantu* by Vyasa Keshavaram Indian Drug Research Institute Poona (1962)
52. A Namjoshi In: *Realms of Ayurveda* (Ed: S Sharma) Arnold-Heinemann New Delhi (1979) p 217
53. P V Sharma *Introduction to Darvyaguna* Chaukhambha Orientalia Varanasi (1976)
54. R D Lele In: *Ayurveda and Modern Medicine* Bharatiya Vidya Bhavan Bombay (1986) p 3

55. B V Subbarayappa In: *India's Contribution to World Thought and Culture* (Eds: L Chandra *et al*) Vivekanand Rock Memorial Committee Madras (1970) p 54
56. J Filiozat In: *India's Contribution to World Thought and Culture* (Eds: L Chandra *et al*) Vivekanand Rock Memorial Committee, Madras (1970) p 67
57. R N Chopra I C Chopra K L Handa and L D Kapur *Chopra's Indigenous Drugs of India* U S Dhur Calcutta (1958)
58. B S Aswal D S Bhakuni A K Goel K Kar and B N Mehrotra *Indian J Experiment Biol* 22 (1984) 487 and earlier references cited therein
59. H K Desai *et al*, *Indian J Chem* 11 (1973) I and earlier references cited therein
60. R P Rastogi and B N Dhawan *Indian J Med Res* 76(Suppl Dec 1982) 27
61. *ICMR Bulletin* 2(4) (1972) 1
62. G V Satyavati M K Raina and M Sharma (Eds): *Medicinal plants of India* 1 Indian Council of Medical Research New Delhi (1976)
63. G V Satyavati A K Gupta and N Tandon (Eds): *Medicinal plants of India* 2 Indian Council of Medical Research New Delhi (1987)
64. M Sittig *Pharmaceutical Manufacturing Encyclopaedia* Noyes Data Corporation New Jersey (1979) p 556
65. RJ Wakil *Br Heart J* 11 (1949) 350
66. Sukh Dev *Science Age* (July 1987) 13
67. J S Tandon M M Dhar S Ramkumar and K Venkatesan *Indian J Chem* 15B (1977) 880
68. A K Saksena *et al Tetrahedron Lett* 26 (1985) 551
69. S V Bhatt B S Bajwa H Dornauer N J De Souza and H W Fehlhaber *Tetrahedron Lett* (1977) 1669
70. M P Dubey R C Srimal and B N Dhawan *Indian J Pharmacol* 6 (1971) 15
71. V P Arya *Drugs of Future* 4(1979) 26
72. K B Seamon *Ann Rep med Chem* 19 (1984) 293
73. *Wealth of India* (Raw Materials) II Council of Scientific and Industrial Research Delhi (1950) p 308
74. Sukh Dev *Arogya-J Health Sci* 3 (1977) 121
75. M L Dhar *et al. Indian J experiment biol* 11 (1973) 43
76. Comments of Dr C Dwarakanath on the views expressed by Dr M L Dhar on the working of CDRS (Circular from Ministry of Health and Family Planning, dated Nov 25 1966)



77. *Wealth of India* (Raw Materials) II, Council of Scientific and Industrial Research Delhi (1950) p 313
78. V V S Sastry *Bull Indian Inst Hist Med* 6 (1976) 102
79. M L Gujral K Sareen K K Tangri M K P Amma and A K Roy *Indian J Physiol Pharmacol* 4 (1960) 267
80. G Santhakumari M L Gujral and K Sareen *Indian J Physiol Pharmacol* 8 (1964) 36
81. G V Satyavati *M D Thesis* Banaras Hindu Univ Varanasi (1966)
82. G V Satyavati C Dwarakanath and S N Tripathi *Indian J med Res* 57 (1969) 1950
83. V V S Shastri *M D Thesis* Banaras Hindu Univ Varanasi (1967)
84. S N Tripathy V V S Shastri and G V Satyavati *J Res Indian Med* 2(2) (1968) 10
85. V L Mehta C L Malhotra and N N Kalra *Indian J Physiol Pharmacol* 12 (1968) 87
86. D S Khanna O P Agarwal S K Gupta and R B Arora *Indian J med Res* 57 (1969) 900
87. V D Patil U R Nayak and Sukh Dev *Components of Guggulu resin from Commiphora mukul* paper presented at Seminar on Disorders of Lipid Metabolism N Delhi Oct 15, 16 (1979)
88. N K Kapoor and S Nitya Anand *Hypocholesterolemic effect of the fractions isolated from C mukul* Paper presented at Seminar on disorders of Lipid Metabolism New Delhi Oct 15, 16 (1971)
89. S Nitya Anand N K Kapoor and Sukh Dev *Proceedings of Indo-German Symposium on Medicinal Plants with Cardiovascular Activity* CDRI Lucknow Feb 22-26 (1977)
90. S Nitya Anand and N K Kapoor *Indian J experiment Biol* 9 (1971) 376; 11 (1973) 395
91. Sukh Dev V D Patil and U R Nayak *Indian Patent* 126354 (1970)
92. (a) V D Patil U R Nayak and Sukh Dev *Tetrahedron* 28 (1972) 2341; (b) 29(1973) 341; (c) 29(1973) 1595
93. R S Prasad and Sukh Dev *Tetrahedron* 32(1976) 1437
94. A G Bajaj and Sukh Dev *Tetrahedron* 38 (1982) 2949
95. For a summary see: *ICMR Bull* 17(1) (1987); K Raghunathan and R Mitra (Eds) *Pharmacognosy of Indigenous Drugs* 1 Central Council for Research in Ayurveda and Siddha New Delhi (1982) p 354
96. S Bose and K C Gupta *Indian J Chem* 2(1964) 57, 156; 4 (1966) 87
97. R P Singh H N Subbarao and Sukh Dev *Tetrahedron* 37(1981) 843

98. H Schildknecht *Angew Chem int Edn* **9** (1970) 1
99. W R Benn and R M Dodson *J Org Chem* **29** (1964) 1142
100. A G Bajaj Sukh Dev E Arnold B Tagle and J Clardy *Tetrahedron Lett* **22**(1981) 4623
101. R S Prasad *Ph.D Thesis* Poona University (1976)
102. A J Birch W V Brown J E T Corrie and B P Moore *J Chem Soc Perkin I* (1972) 2653
103. A G Bajaj *Ph.D Thesis* M S Univ Baroda (1980)
104. Vijay Kumar *Ph D Thesis* M S Univ Baroda (1984)
105. Vijay Kumar and Sukh Dev *Tetrahedron* **43** (1987) 5933
106. R H Gigg In: *Rodd's Chemistry of Carbon Compounds* (Ed : S Coffey) **IE** Elsevier, Amsterdam (1976) p 397
107. *Gugulipid, Phase III Clinical Data and Summary of Preclinical and Phase I and II Clinical Data* Central Drug Research Institute Lucknow (1986)
108. U K Kanitkar P S Dange and G S Pendse *J Res Indian Med* **3** (1969) 123
109. A K Nadkarni *Indian Materia Medica I* Popular Book Depot Bombay (1954) p 153
110. P V Sharma *Dravyaguna-Vijnana II* Chaukhamba Bharti Academy Varanasi (1986)
111. B B Gaitonde and M H Jethmalani *Arch int Pharmacodyn* **179** (1969) 121
112. P B Sabnis B B Gaitonde and M H Jethmalani *Indian J experiment Biol* **6** (1968) 55
113. A B Patel and U K Kanitkar *Indian Vet J* **46** (1969) 718
114. P R Ravikumar R Soman G L Chetty R C Pandey and Sukh Dev *Indian J Chem* **26B** (1987) 1012
115. J Joshi and Sukh Dev *Indian J Chem* **27B** (1988) 12.
116. See eg (a) Various issues of *J Res Indian Med*, *J Res Ayur Siddha*, *Indian J med Res*, *Indian J experiment Biol* (b) Refs 62, 63
117. S S Gupta S C L Verma A K Ram and R M Tripathi *Indian J Pharmacol* **4** (1972) 208
118. C K Atal *et al Chemistry and Pharmacology of Vasicine, A New Oxytocic and Abortifacient* Regional Research Laboratory Jammu (1980)
119. N S Mooss *Ayurvedic Flora Medica* Vaidyasathya Press Kottayam (1978)
120. C D Dey S Bose and S Mitra *India J Physiol alld Sci* **30** (1976) 88
121. Anon *Med Arom Plant Abstr* **3**(6) (1981) 75

122. J Lal S Chandra V Raviprakash and M Sabir *Indian J Physiol Pharmacol* **20** (1976) 64
123. K Kar *et al J Pharmac Sci* **64** (1975) 258
124. K S Karanth T K Padma and M N Guruswami *Arogya-J Health Sci* **7**(1) (1981) 83
125. Anon *Med Arom Plant Abstr* **8** (1986) 386
126. Anon *Med Arom Plant Abstr* **4** (1982) 351
127. H Wagner In: *Natural Products as Medicinal Agents* (Eds. J L Beal and E Reinhard) Hippokrates Verlag Stuttgart (1981) p 217
128. O P Gupta C K Atal and J G Langar *Indian J Pharmacol* **13** (1981) 60
129. A Prakash and R K Verma *Bull med-ethn-bot Res* **2** (1981) 266
130. H Goyal S Sukumar and K K Purushothaman *J Res Ayur Siddha* **2**(1981) 286
131. K D Sharma M L Kapoor S P Vaidya and L K Sharma *J Res Ayur Siddha* **2**(1981) 309
132. Anon *Med Arom Plant Abst* **5** (1983) 273
133. S Venkataraghavan *et al J Res Ayur Siddha* **1**(1980) 370
134. V N Pandey *J Res Ayur Siddha* **2**(1981) 178
135. H König *Angew Chem int Edn* **19** (1980) 749



**Chintamani Nagesa Ramachandra Rao** (b. 30 June 1934) did Ph.D. (1958) from Purdue University, USA; D.Sc. (1960) from University of Mysore and received D.Sc (h.c.) from 22 universities. He is Director, Indian Institute of Science; and various faculty positions at IISc and Indian Institute of Technology, Kanpur.

Rao has contributed widely to several aspects of spectroscopy and structure by employing vibrational, electronic and high-energy spectroscopies. Besides complex molecules, he has investigated molecular interactions by these techniques. An important area

of vital interest to which he has made monumental contributions is solid state chemistry. In this area, he has carried out extensive studies related to tailoring of complex solids, structure-property relations, defect solids, phase transitions, superconducting materials and novel synthetic methods. Also contributed significantly to surface science, fullerenes and metal clusters.

Rao is the Founder President, Materials Research Society of India and Indian Academy of Sciences; Fellow, The Royal Society, London; Foreign Associate, National Academy of Sciences (USA); Honorary Foreign Member, American Academy of Arts & Sciences; Foreign Member, Russian Academy of Sciences, Polish Academy of Sciences, Czechoslovakian Academy of Sciences, Serbian Academy of Sciences and Slovenian Academy of Sciences; Member, Pontifical Academy of Sciences; Founder Fellow (also Vice-President), Third World Academy of Sciences; Honorary Fellow, Royal Society of Chemistry (London), Institution of Engineers (India) and Institution of Electronics and Telecommunication Engineers (India); Honorary Member, Materials Research Societies of Japan and Korea, and International Academy of Ceramics; President, INSA (1985-86). President, Indian Science Congress Association; Centennial Foreign Fellowship (American Chemical Society) (1976); FICCI Award (1977); S.N. Bose Medal (INSA) (1980); The Royal Society of Chemistry (London) Medal (1981); P.C. Ray Medal (Indian Chemical Society) (1984); Jawaharlal Nehru Award (1988); G.M. Modi Award (1989); Hevrovsky Gold Medal (Czechoslovak Academy of Sciences) (1989); Meghnad Saha Medal (INSA) (1990); CSIR Golden Jubilee Award (1991); Kamal Kumari Foundation Award (1992); Padma Vibhushan (1985).

*Chintamani Nagesa Ramachandra Rao was elected to the fellowship of the Academy in 1974.*

**Address:** Director, Indian Institute of Science and , President, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 012; Tel: (080) 341690 (Off); 340580 (Lab); 341254, 345910 (Res); Telex: 0845-8349 IISC IN; Fax: 91-80-341683; E-mail: iiscamp@vigyan.ernet.in.

## DIMENSIONS OF MATERIALS RESEARCH

C N R RAO FNA, FRS

*The varied dimensions of materials research are examined in relation to India in the context of the world scenario. The way materials science works today is illustrated by chosen real-life examples. The problems and challenges posed by the new superconducting cuprates are presented as a case study. The author's interest in the area of oxide materials, specially superconductors, are also briefly indicated. It is pointed out how in this highly competitive area, it is essential to develop the right strategies in order to find a meaningful role for ourselves.*

### PROLOGUE

I consider the award of a General Medal of the Academy to a Fellow to imply that the significance and magnitude of the contributions of the awardee to science are more than satisfactory. I am flattered to be considered worthy of this General Medal and I offer my sincere thanks to the President and the Council of the Academy. I feel specially honoured since this medal is named after a truly great Indian Scientist, Meghnad Saha.

Following the good traditions of the Academy, I shall take a scientific theme of vital significance for the award lecture. I have chosen materials research as the topic since the progress of mankind has been closely linked to man's quest for new and improved materials. It is also a highly competitive area of great national importance. I shall discuss the world scenario as well as new directions in this area and examine some aspects relevant to our development. I shall cite a few typical examples from the wide spectrum of materials to illustrate how materials science works and in doing so, I cannot help being partial to non-metallic materials. A medal lecture is also the time for taking stock of what one has done in one's research career and to account for the support one has received. I shall do this as well, though briefly.

Materials science is distinctly different from physics and chemistry of materials. Physics and chemistry get transformed into materials science wherever they are motivated by an engineering objective. Materials science can be represented by a tetrahedron (Fig. 1) having as its base,

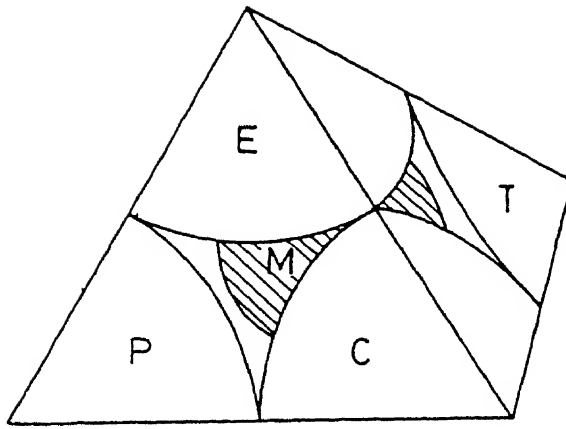


Fig. 1 Materials Science (M) at the centre of a tetrahedra: P, Physics; C, Chemistry; T, Theory; E, Engineering Design.

fundamental sciences such as physics and chemistry, embracing both experimental and theoretical aspects. The tetrahedron is completed by the engineering design associated with a technological target. At the centre, holding the disparate disciplines is the interdisciplinary area of materials science. Engineering design requires materials specifications which are generally formulated in terms of a set of engineering parameters that define the functions to be performed. In order to optimize the engineering parameters, the engineer requires a theoretical framework to translate from the language of design to the variables or the phenomenological parameters such as shape, size and history. If, however, a material is to be designed to satisfy specific engineering requirements, the theoretical framework should include not only the connections between engineering and the phenomenological parameters, but also a proper understanding of how chemical composition, structure and bonding of the material are related to the particular properties of interest. Without an understanding of these relationships, material design aimed at performing a new or an improved function would be left to empiricism. Over the last few years, our fundamental understanding of materials has become sufficiently advanced and sophisticated to make a demand on those in charge of public funds for science and technology to specify and prioritize the engineering and technological targets sufficiently so that concentrated support to materials R&D can lead to substantial public benefit. Today, I do not plan to provide a history of selected examples to illustrate how materials

science has led to such revolutionary technologies as digital computers, lasers, fibre optics, plastics or composites. Rather, I shall attempt to show how the mind of a materials scientist seems to function in his search for new directions and in his thirst for new or alternative materials in the competitive world of today. In making such a presentation, I shall attempt to bring in the Indian focus in the context of world trends, besides some of my own research interests.

### THE WORLD SCENARIO

Today's world is propelled by man's urge to live in comfort without excessive dependence on conventional manual labour and to surpass himself in communications, information and transportation. We have to accept this to be the situation in the advanced countries without arguing whether such a life is desirable or necessary. It is also true that most of the developing countries are nowhere near having such a situation, but it would be unrealistic to assume that they are not going to be affected by the trends in the advanced countries. Whether one likes it or not, the poorer countries will eventually buy or borrow products of modern technology from the advanced countries.

Much of the progress of mankind today is directly or indirectly dependent on advanced technology materials which perform better and with new dimensions. For some time, most of the emphasis was on metallic materials. Slowly, metallic materials are getting replaced by non-metallic ones even for use as structural materials. Today, composites and polymeric materials occupy a prime position as structural materials. Innovations in non-structural materials have been even more stupendous. New types of sensors have been devised for every possible situation. Ceramic materials have become a major component in electronic, electrical, automotive and other industries. Man is excelling himself in making better and faster "chips".

Excessive use of carbon-containing polymers (mostly based on petrochemicals), much of its progressively replacing metals, raises the question as to what man will do when he has exhausted most, if not all, of the natural carbon resources such as petroleum. This could happen by the middle or the end of the 21st century. It is possible that man will then make use of inorganic carbon sources (carbonate rocks) to make hydrocarbons and polymers.

Let us look at the key advanced technology materials today. These are:

- Composites (Structural materials)
- Materials for electronics and photonics
- Ceramic materials for a variety of uses

The key sectors are:

- Communications (ferrites, fibres, lasers)
- Energy (New batteries, fuel cells, solar energy beneficiation)
- Transportation (composites, superconductors)

In these sectors, major revolutions may occur in the next few years. For example, amorphous silicon may soon emerge to be efficient for solar photovoltaic applications. High-temperature superconductors may revolutionize surface transportation (e.g., levitating trains). Even in conventional areas, new materials are likely to outstrip the old ones in performance. For example, new magnet materials (e.g. Nd-Fe-B, nitrided rare earth alloys) are so much better than the old ferrites or metallic systems. Photonics is slowly replacing electronics in communications and information technology. There has also been progress in biomaterials (for replacements, transplants etc.)

### THE MIND OF THE MATERIALS SCIENTIST

The rapid pace with which the materials world has progressed is clearly due to the uncanny knack of the materials scientist who does not leave any stone unturned in his search for new and better materials. This effort is really mindboggling. The following are some of the factors that have been responsible for the advances made:

- Study of known or even common materials (both natural and unnatural)
- Serendipity
- Design (tailor making and engineering of materials)
- Creation of demand for new materials

All these characteristics would be obvious to everyone. I would like to underscore, however, the need to carefully look at common as well as naturally occurring materials. Coral, which occurs in nature, has given new ideas for designing ferroic composites. Something as common as fish



scale shows interesting properties (e.g., high dielectric constant as shown in Fig. 2) The last factor requiring materials scientists to create a demand (or a market) for their own area is of some significance; in fact, this is true of all of science and technology as far as India is concerned.

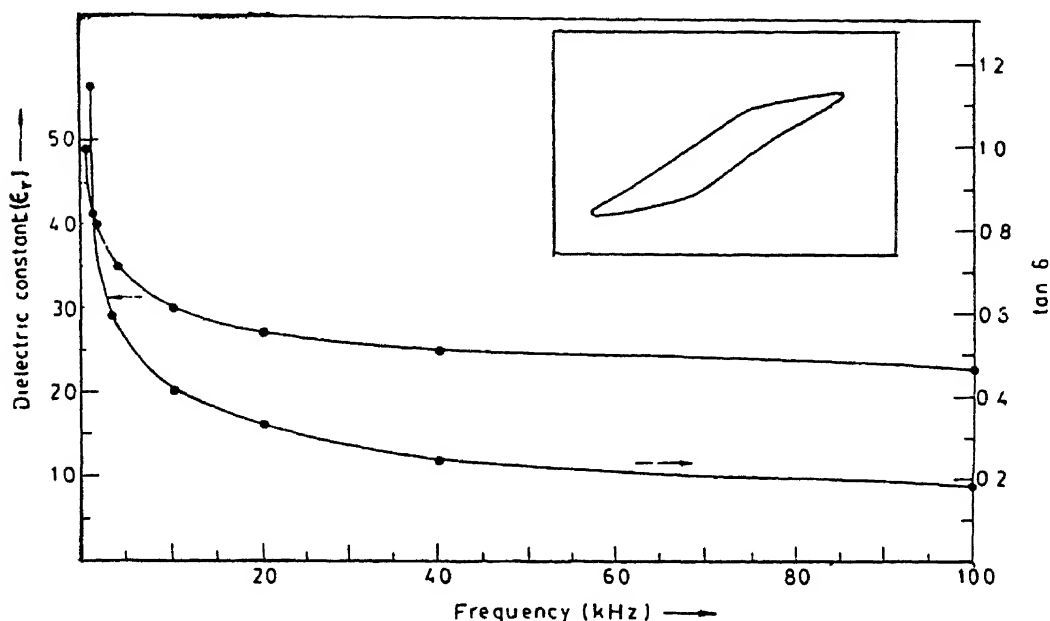


FIG 2 Dielectric constant of common fish scale. Note the relatively high values and also the dielectric hysteresis given in the inset (Work K B R Varma of this laboratory).

Many real-life examples can be given for each of the above aspects of materials science. Some of the typical examples that I would cite are the following:

- Silicon nitrides (Si-Al-O-N)
- Diamond films
- Zeolites and clays (controlled pore inorganic solids)
- The Lanxide process (new composites)
- New sensors (using oxide and other materials including microbiosensors).

I cannot but admire the ingenuity of Ken Jack when he decided to replace part of the silicon and nitrogen in silicon nitride (Si-N) by aluminium and oxygen respectively to make Si-Al-O-N (Sialon) expecting it to be an improved high temperature ceramic material. I recall how Deriyagin carried out the first experiments to break hydrocarbons

such as ethylene and methane in a reducing atmosphere to deposit carbon in the form of diamond on a metal surface. Diamond films are being investigated in a number of laboratories today and they are likely to find many applications. Zeolites occur in nature, but chemists have synthesised a variety of zeolites which permit shape-selective catalysis by virtue of the specificity of the cage sizes in them. Clays which possess sheet structures can similarly be modified to carry out selective reactions. A most unusual example of serendipity and doggedness is provided by the Lanxide process. Newkirk accidentally found sometime ago that the stuff that collects on the top of molten aluminium is not a useless muck, but an unusual composite containing the metal and aluminium oxide. Processes have since been developed to make this material in large quantities to produce fibres with optimal mechanical properties. Over 200 patents have been filed in the last three years. We still do not know what real uses will be found for the unusual composites obtained by the Lanxide process, but a company has been set up solely based on this innovation.

Sensors clearly provide the finest examples of the versatility and pragmatism of materials scientists. Today, we have sensors for most gases and vapours, for heat and for other forms of radiation. The sensor industry is booming, with Japan taking the lion's share of the market. Most sensors use simple materials (e.g. oxides such as  $\text{ZrO}_2$ ,  $\text{SnO}_2$  and  $\text{ZnO}$ ) and a device to measure the signal (usually, electrical resistance). The sensor industry is a knowledge-based industry with high returns for a marginal investment. The Japanese are now working on a variety of biosensors which will make use of specific biological systems (e.g., enzymes) along with appropriate electrodes and measuring devices. For example, a needle-shaped sensor can rightaway tell a customer in the supermarket whether the pineapple or melon is ripe or not. I can imagine a new world with people carrying not only the needle shaped sensors to buy fruits, but a variety of other sensors related to human health (e.g., for diabetes, cholesterol level), environment and even to determine the cleanliness of toilets.

What I have listed above is at best representative and there are probably better examples one can cite. There is much science behind all such efforts and this is often forgotten. Newer and improved methods of synthesis and characterization are the essential aspects of this science. Materials scientists are becoming fearless and are taking on more and more complex materials in all states of aggregation for possible exploitation.

Many fine examples can be given to illustrate the art and science of materials synthesis. Novel, less energy-consuming methods of synthesis are being developed all the time. To give a chemical example,<sup>1</sup> the precursor route has eliminated the limitations of the old brute-force methods to yield homogenous pure substances (e.g., synthesis of Chevrel phases and many complex oxides by the precursor route). Positive and negative pressure effects are attained by appropriate chemical substitution (e.g., partial substitution of V by Ti or Cr in  $V_2O_3$ ). Similarly, insulators are transformed into metals or superconductors by composition control. The basic tenet is that there must be a way of making any given material in any required shape or form. High-pressure, ultra-high vacuum, very high or low temperatures, plasma, rapid cooling and a variety of other means have been employed to attain the desired conditions for synthesis. A trend worth noting is that many so-called metastable materials are prepared in this way for use under ordinary conditions. Metastability is part of the innovation in materials.

Techniques of materials characterization have undergone a dramatic change in the last few years<sup>1</sup>. The power of analysis (composition, structure or both) on the tiniest of samples (a few atoms) available today is truly remarkable. The advent of atomic force microscopy and scanning tunneling microscopy enables the study of materials at atomic level under real conditions (in air, with liquid interface, in vacuum etc.)

### THE INDIAN SCENARIO

The importance of materials for India was pointed out as early as in 1971 by the first National Committee on Science and Technology (NCST) of which I had the privilege of being a member. The NCST produced a report<sup>2</sup> on Indian needs for special materials. Recently, the erstwhile Science Advisory Council to the Prime Minister has produced an excellent document<sup>3</sup> on the materials needs of India in the world context. We, in India, are in a unique position where we have to worry about materials for basic needs as well as for advanced technology. Typical materials related to our basic needs would be:

- Housing (Alternative, inexpensive construction materials; replacement for wood which is scarce; minimization of the use of cement)
- Food (packaging materials)
- Health (biomaterials for use in hospitals and in patients)

Some aspects of the Indian scenario that we need to note are the following:

- Large quantities required of all materials because of our large requirements.
- Availability of the science and technology base and therefore the possibility of using the knowledge of materials science to solve problems of society.
- Unlimited market and future for materials in various sectors, some with export potential.

We have little choice but to develop our own base in advanced technology materials. Our materials industry is at best in its infancy and we have to grow. If we have to solve our serious economic problems, we would have to develop knowledge-based industries many of which would depend on key materials. We have to export value-added products and reach an export figure of at least Rs. 100,000 crores per annum in the next 5 years. This can be done if a judicious materials policy is developed wherein we decide clearly what to make, what to sell and what to import. There are many natural resources that we need to properly exploit. We have to decide on what and how much of mineral assets to sell and hoard. Typical of them are the monazite sand and the titanium ore. We still use imported nickel-based alloys in many of our installations instead of titanium. Even though expensive, there may be advantage in using our own indigenous materials. Our buildings over-use expensive construction materials. We have not found enough use for what we make in plenty (e.g., aluminium). We have not learnt to make standard materials (e.g., steel) at a reasonable cost. We have got used to a culture which is against innovation; some industries do not even have the all-important profit-motive. In addition, we are not clear-headed about the most essential things that we need.

Let us take the example of photovoltaics. We still have not decided as to how much of solar photovoltaic power we will produce indigenously for the rural sector. Will it be 25 or 100MW in the 8th Plan? In the 7th Plan, it was supposed to be 25 MW, but we just ended up with less than 5MW. Depending on this target, we have to produce the required polycrystalline silicon which is then to be transformed into single crystals and then to wafers. Roughly 25 tonnes of polycrystalline silicon are required for 1MW of solar photovoltaic power based on crystalline silicon

wafer cells. Today, we are just able to produce around 25 tonnes per annum of polycrystalline silicon at Mettur. Since the strategy for photovoltaics for the next 5 years will have to be based on crystalline silicon wafers, it is important that we have an exact short-range target. In the long term, it is possible that amorphous silicon films may replace the crystalline wafers. We would then not require as much polycrystalline silicon. The development of our silicon industry, therefore, crucially hinges on our national priorities as much as on new technologies.

The report of the Science Advisory Council to the Prime Minister<sup>3</sup> lists the following key materials to be crucial to India's development:

- **Electronics Sector**  
Silicon (crystalline/amorphous).  
Gallium arsenide and other semiconductors  
Superconducting oxides
- **Communications Sector**  
Fibres  
Semiconductors (phosphors, lasers)  
Magnetic materials (Fe-Nd-B, recording materials)  
Non-linear optic materials (including both inorganic and organic)
- **Transportation Sector**  
Aluminium alloys  
Toughened zirconia  
Kevlar and carbon-carbon composites
- **Energy sector**  
Amorphous silicon (photovoltaics)  
Hydrogen storage materials  
New batteries
- **Other industrial sectors**  
Sensors (oxides)  
Lasers (for machining, annealing etc)  
Superconductors (for magnets)  
Concentrated and comprehensive effort is recommended with regard to the following advanced materials:

- |                                |   |
|--------------------------------|---|
| ● <i>Metals and Alloys:</i>    | Titanium  |
| ● <i>Electronic Materials:</i> | Silicon based micro-electronics and<br>III-V semiconductors (Ga As) |
| ● <i>Magnetic Materials:</i>   | Nd-Fe-B, Superconducting Nb alloys                                  |
| ● <i>Ceramics</i>              | Zirconia  |

● *Composites*

- (a) Carbon, SiC, Kevlar fibres
- (b) Carbon composites

It is hightime that we become materials-conscious and develop a well-orchestrated policy which will pay proper attention to all aspects of the innovation chain. Since India cannot afford to be behind economically and industrially, there is little choice for us in this regard. We have to recognize our needs and challenges in materials and start our action plan. Given the right support and direction, nothing should stop us from reaching where we want to go. At the same time, we have to be conscious of newer developments so that when the time comes, we can join the race. If we do not, we shall soon be having a new colonial era when we will owe our existence to the advanced countries who will be able to demand the price they choose for materials crucial for our development.

### SUPERCONDUCTORS—A CASE STUDY

In order to illustrate how a typical area of materials grows with time, I shall take the example of superconducting materials. Superconductivity has been a fascinating subject since the discovery by Kammerlingh Onnes in 1911 that mercury becomes superconducting at 4.2K. The superconducting state of a solid when it offers no resistance to the passage of electricity and repels a magnet has not only been a favourite ground for scientists interested in the study of solids, but also of immense interest to engineers and technologists because of the tremendous applications. We have been beneficiaries of superconductor technology for some time. NMR imaging employed in medical diagnostics instead of the traditional catscan employing X-rays uses large superconducting magnets. Many electronic devices using superconducting materials have been fabricated. Although power transmission or energy storage using superconducting wires has not become a reality, there is every likelihood that there will soon be high-speed levitating trains using superconductor technology. All this has been done with materials which become superconducting at liquid helium temperatures. The discovery of the new copper oxide superconductors ( $T_c \sim 30$  K) by Bednorz and Müller<sup>4</sup> in 1986 opened up new exciting possibilities. Today, we have many ceramic oxide materials which become superconducting well above liquid nitrogen temperature<sup>5</sup> (Fig. 3.) There is every chance that some day one may discover materials which become superconducting even at higher temperatures, possibly close to room temperature. It would be truly remarkable if we can have

large superconducting magnets which at best would require chilled water for cooling.

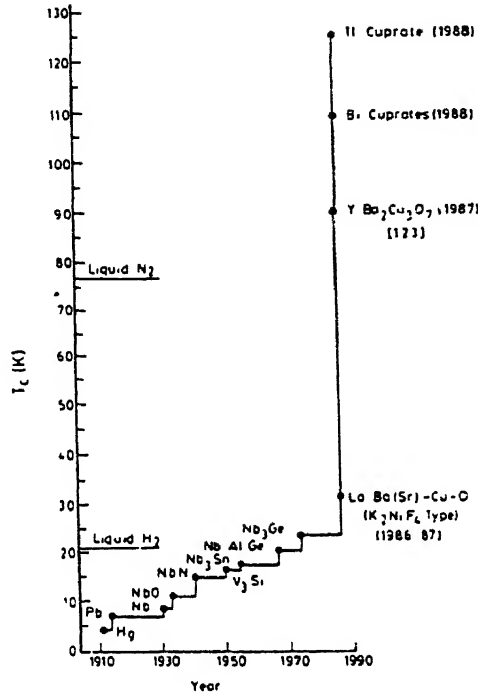


FIG 3 Superconductivity over the years. The transition temperature is plotted against the year.

The challenges and opportunities offered by high temperature superconductors are really exciting because of their scientific importance as well as technological value. Japan, USA and many other countries are competing intensely for leadership in this field. Besides increased funding for scientific research, a variety of efforts are being made in areas related to technology. Consortia have been formed by industries and new modes of collaboration have been initiated between educational, industrial and governmental organizations.

Some of the important objectives of research and development programmes in superconductivity today are:

- (i) to acquire better and additional experimental data on the existing high  $T_c$  materials in order to improve our understanding of the essential properties:

*(This is important since only measurements made on pure, homogenous materials are of real significance. This is specially true of measurements made on relatively large single crystals).*

- (ii) to develop suitable models to understand the basic mechanism responsible for high-temperature superconductivity and related aspects of these materials;

*(The well-known BCS model does not seem to describe the new oxide superconductors. Although there has been a spate of theoretical papers on high-temperature superconductivity of oxide materials, there is yet no model which is completely satisfactory. It is necessary to have models which interpret known results and also make predictions that can be verified).*

- (iii) to look for new materials (with or without copper) exhibiting high  $T_c$ ;

*(It is essential that the search for newer materials with high  $T_c$  continues. One hope is that someday one will find a material with  $T_c$  close to room temperature. Another important effort should be to look for non-copper materials with high  $T_c$ . If we can discover them, the situation with theory will change dramatically since most models require the d-orbitals of Cu).*

- (iv) to improve our understanding of the chemistry; process parameters and of the ceramic properties of oxide superconductors;

*(This is an area of great importance since we do not know enough about the processing of ceramic powders. New methods of synthesis and continuous processing techniques need to be worked on thoroughly).*

- (v) to prepare good quality high  $T_c$  films by a variety of techniques (plasma, laser ablation, rf-sputtering, electron beam evaporation, MOCVD etc.) for electronic device fabrication;

*(Most of the superconductor applications in electronics require the materials to be in the form of films. It is therefore necessary to make homogeneous films with good current carrying capacity).*

- (vi) To prepare high  $T_c$  bulk superconductors in the form of tapes and wires with high current carrying capacity;

*(The main limitation in the use of oxide superconductors for high-field, large-scale applications is that we cannot make wires or tapes of these materials with high current density. The difficulty is not because of the ceramic nature of the materials alone, but also with the inherently low current density of these granular materials. If by*



*proper orientation of particles and other techniques, we can get high  $T_c$  tapes or wires which can operate in moderate to large magnetic fields, then the applications of these materials would be many. There has been some success in this direction in Japan).*

- (vii) to fabricate SQUIDS, detectors, superconductor-semiconductor hybrids and other electronic devices for various applications with high  $T_c$  materials;

*(There has already been some success in this direction, but more effort is necessary to get reliable, commercially viable devices).*

- (viii) to build small proto-type generators, magnets etc with ceramic materials, in order to gain experience;

*(This experience would be necessary if and when we have the right ceramic materials with high  $T_c$ ).*

- (ix) To build prototype levitating trains and continue to experiment on shipdrive systems; and

*(The use of superconductors in designing levitating trains is limited to Japan. More effort is necessary in this area. Experimental trains using ceramic tapes may be worthwhile. Experience already gained in ship-drive systems has to be consolidated).*

- (x) in India, there is not adequate expertise in building large magnets even with the conventional liquid helium superconductors. It would be good to make a beginning now, even though late. This could help, for example, in fabricating magnets for NMR imaging required in large numbers in hospitals. It is to be noted that based on the recommendation of NCST, BHEL was supposed to take up some major projects using Nb alloys as early as in 1973, but this did not happen.

### SOME ASPECTS OF PERSONAL RESEARCH

No talk in the Academy by a Fellow should end without some reference to the Fellow's own research work conducted in the recent past. I shall therefore briefly mention one line of research related to materials that I have been involved in for over three decades.

Transition metal oxides have constituted one of my main areas of research for several years.<sup>1,6</sup> These oxides exhibit a whole gamut of properties because the metal-oxygen bond is not too ionic as the metal

fluorine bond or too covalent as the metal-sulphur-bond. This is why metal oxides exhibit a range of properties with insulating behaviour at one end and metallic or superconducting behaviour at the other. My students and co-workers have synthesised and characterized a variety of oxides, more importantly, have designed several novel oxides with the desired properties. Many of these oxides have found use as catalysts, electronic ceramics and so on. I am specially happy that the new superconductor revolution has been based on layered oxides of copper, a class of materials which my co-workers and I have been investigating for many years. Since the discovery of high  $T_c$  superconductivity in these materials, we have been engaged actively not only in the synthesis and characterization of novel superconducting oxide materials, but also in carrying out experiments designed to understand some aspects of the mechanism of superconductivity in cuprates.<sup>6-9</sup> Let me list some of the important contributions that we have made in this field before I make a brief mention of some of the work that we are pursuing at present.

*Some highlights of oxide research of the author  
related to superconductivity*

| <i>Particulars</i>  | <i>First Publication</i>                                |
|---|---|
| ● Electronic properties of Perovskite oxides  | 1969, 1971, <i>J Phys Chem Solids, Phys Rev</i>         |
| ● Structure and properties of layered oxides of $K_2NiF_4$ structure                | 1973, <i>Mater Res Bull</i>                             |
| ● An overview of layered oxides   | 1984, <i>J Solid State Chem</i>                         |
| ● Independent identification of $YBa_2Cu_3O_{7-x}$ (90K superconductor)             | 1987, <i>Nature</i>                                     |
| ● First observation of non-resonant microwave absorption by cuprate superconductors | 1987, <i>J Phys C Solid State</i>                       |
| ● Importance of oxygen holes and hole pairing in the mechanism of superconductivity | 1987, 1988, <i>Phys Rev, Physica C</i>                  |
| ● Studies of bismuth and thallium cuprate superconductors                           | 1988, <i>Physica C, Solid State Commun</i>              |
| ● New Families of thallium and cuprate superconductors                              | 1988, 1989, 1990 <i>Phys bismuth Rev, App Phys Lett</i> |
| ● Possible occurrence of superconductivity in nickelates                            | 1989, <i>Solid State Commun</i>                         |
| ● Metal oxidation states in oxide superconductors                                   | 1989, 1990 <i>App Phys Lett, Phys Rev</i>               |
| ● Non-existence of the so-called 60 K superconducting phase in 123 cuprates         | 1990 <i>Phys Rev</i>                                    |
| ● Importance of charge-transfer energy to superconductivity                         |   |

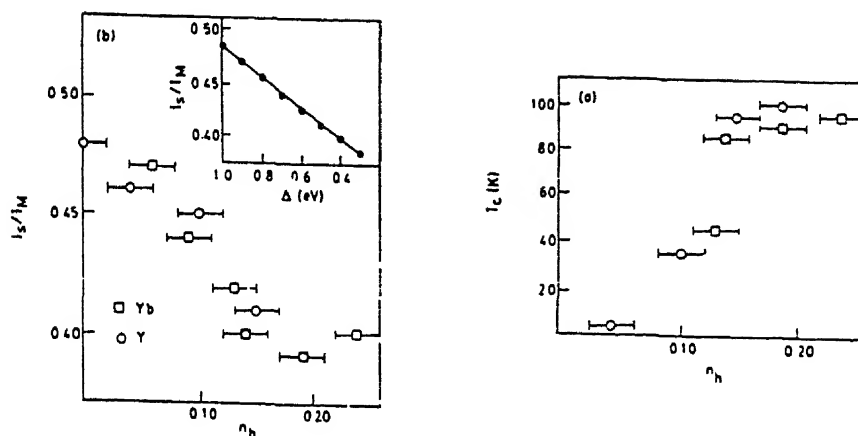


FIG 4 (a) Variation of  $T_c$  with the number of holes in bismuth cuprates.  
 (b) Variation of the Cu(2p) relative satellite intensity,  $I_s/I_M$  with the number of holes in bismuth cuprates. Inset shows the dependence of the satellite intensity on the charge-transfer energy

The search for new cuprate superconductors continues. In addition, in the last few months we have been looking for copper-free oxides which could show superconductivity. This is important since most of the recent models for superconductivity require the  $d$ -electrons of copper. Encouraged by the finding<sup>10</sup> that the Cu-O charge-transfer excitation energy and the  $p$ - $d$  hybridization strength which determine the intensity of the Cu(2p) satellite in X-ray photoemission, also play a crucial role in the superconductivity of the cuprates (Fig.4), we are exploring several other systems experimentally to look at this aspect more carefully. At the same time, we are examining to see how the polarizability of oxygen is affected by the  $p$ - $d$  hybridization strength and the charge-transfer energy. Clearly, polarization will increase with a decrease in the charge-transfer energy, the latter also favouring the formation of oxygen holes. We believe that this is a useful direction to follow since the polarizability of oxygen has been known to vary widely from one oxide to another. Depending on the proportion of the  $O^{1-}$ (hole) type state relative to the  $O^{2-}$  state (ionic limit), we can have a range of oxygen polarizabilities in metal oxides. It is not unlikely that polarizabilities closer to the ionic limit prevail in ferroelectrics while those closer to the  $O^1$  limit are relevant cuprate superconductors.

## EPILOGUE

It is not unlikely that ceramic materials will be used in turbines and automobile engines in the near future. It is quite likely that the photon era will replace the electron era and we will soon have very efficient optical communication and optical computers. New aeroplanes will be made almost entirely of composites, as also a good part of the automobiles. Levitating trains will probably operate in many countries. The future looks not only unlimited, but also scary. Today's trends are such that "*who will dominate the world*" will be determined by "*who will dominate advanced materials science and technology*". Will it be Japan, Germany or the U.S.? Clearly, technological leadership will be the requisite for international leadership. Yet, in spite of such a world scenario, except probably in Japan, there has been near absence (or, at the very best, lukewarm response) by Governments to urgent needs; this is not in India alone. The industry's response even in some advanced countries has not been altogether promotive. Market forces alone seem to determine the industry's investment in materials in some of them and this would be bad for the future (an important lesson to be learnt by the advanced countries such as the U.S. as well as a less-developed country like India). We cannot be caught flat-footed when something comes along to change technological trends in vital sectors in such a big way as to change our life-styles altogether.

## ACKNOWLEDGEMENT

The author would like to record his deep sense of appreciation and gratitude to his students and co-workers who, over a period of more than 30 years, have worked with him in the area of oxide materials. The author is thankful to the various supporting agencies, especially the Department of Science and Technology, the University Grants Commission and the US National Science Foundation for support of his research in oxide materials.

## References

1. C N R Rao and J Gopalakrishnan *New Directions in Solid State Chemistry* Cambridge Univ Press (Paperback) (1989)
2. *Special Materials* Rep. of the NCST New Delhi (1973)
3. *Perspectives in Science and Technology* Tech Rep to PM, 1 Har Anand Publications, Vikas Publishing House, New Delhi (1990)

4. J G Bendorz and K Muller *Z Phys B* **64** (1986) 189
5. S K Joshi, C N R Rao and S V Subramanyam (*Eds*) *Superconductivity ICSC Proceedings*, World Scientific Singapore (1990)
6. C N R Rao *Ann Rev Phys Chem* **40** (1989) 291
7. C N R Rao (*Ed.*) *Chemistry of Oxide Superconductors* Blackwell Scientific Publications Oxford (1988)
8. C N R Rao and B Raveau *Accts Chem Res* **22** (1989) 106
9. T V Ramakrishnan and C N R Rao *J Phys Chem* **93** (1989) 4414
10. C N R Rao, G Ranga Rao, M K Rajumon and D D Sarma *Phys Rev B* **42** (1990) 1026



**Sivaramakrishna Chandrasekhar** (b. 6 August 1930) did D.Sc (1955) from Nagpur University and Ph.D.(1958) and Sc.D. (1987), University of Cambridge, UK. He is the Director, Centre for Liquid Crystal Research, Bangalore, Hony. Professor, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore (1991- ), and Bhatnagar Fellow, CSIR (1990- ).

Chandrasekhar's early work was on crystalline optical activity and X-ray diffraction, but he changed his field to liquid crystals in 1961, when he joined University of Mysore. It is generally accepted that the worldwide revival of interest in liquid crystals was due, in no small measure, to the pioneering work of Chandrasekhar and his colleagues. With his versatile interests, he has contributed to many aspects of this subject, but probably his best known work is his prediction and discovery (along with his students B.K. Sadashiva and K.A. Suresh) of discotic liquid crystals. He has authored 'Liquid Crystals' (CUP, 1977, 1992), hailed internationally as a classic; edited three volumes including two on 'Liquid Crystals' (conference proceedings); editor, *Mol. Cryst. Liq. Cryst.* (Gordon & Breach).

Chandrasekhar is the Fellow of Royal Society of London, Institute of Physics (London) and Indian Academy of Sciences; Hony. Fellow, National Academy of Sciences (India); Founding Fellow, Third World Academy of Sciences; Vice President, INSA (1989-90); Founder President, International Liquid Crystal Society (registered in the USA) (1990-92) and Indian Liquid Crystal Society. He is the recipient of UGC National Lecturer (1971-72); CSIR Silver Jubilee Award (1972); Shanti Swarup Bhatnagar Prize (1972); FICCI Award (1979); C.V.Raman Award (UGC) (1981); Mahendra Lal Sircar Award (1984); Karnataka Rajyotsava Award (1986); Homi Bhabha Medal (INSA) (1987); C.V.Raman Centenary Medal (IACS) (1988); Jawaharlal Nehru Award (Madhya Pradesh Council of Science and Technology) (1991); Meghnad Saha Medal (INSA) (1992); R.D.Birla Award (Indian Physics Association) (1992); C.V.Raman Birth Centenary Award (ISCA) (1993).

*Sivaramakrishna Chandrasekhar was elected to the Fellowship of the Academy in 1978.*

**Address:** Director, Centre for Liquid Crystal Research, P.B.No. 1329, Jalahalli, Bangalore 560 013; Tel: (080) 340124 (Ext 226), 337123 (Off), (080) 331989 (Res).

## DISCOTIC LIQUID CRYSTALS

S CHANDRASEKHAR FNA, FRS

*Liquid crystals of disc-shaped molecules, or discotic liquid crystals as they are now called, were discovered just 15 years ago. The subject has since grown enormously. about 1000 discotic compounds are known to date and some 800 papers have been published on the chemistry and physics of these materials. Of late, there is also considerable interest in the field of discotic polymers, a new class of polymer liquid crystals which exhibit novel mesophase structures.*

*The purpose of this talk is to bring out the relevance of molecular engineering in this field of research. The major advances in the area are reviewed briefly, emphasizing the structural aspects rather than their theoretical implications. Some examples are presented of discotic mesogens and discotic polymers, and the structures of the mesophases identified so far are described.*

*As a further illustration of the role of molecular engineering, the design of molecules that led to the significant discovery of the biaxial nematic phase in simple thermotropic systems is also discussed.*

### INTRODUCTION

The vast majority of liquid crystalline compounds are composed of rod-shaped molecules. Indeed, since the early work of Reinitzer, Lehmann and others about 100 years ago, the generally accepted principle was, until fairly recently, that the molecule has to be long and rod-like for thermotropic mesomorphism to occur, but it is now well established that compounds composed of relatively simple disc-shaped molecules may also form stable liquid crystals. The first examples of this kind of mesomorphism were observed in the hexa-substituted esters of benzene, [(Fig. 1(a))]: these compounds were prepared at my instance by my colleague Dr B K Sadashiva (then a student working for his Ph.D). and we found that some homologues did, in fact, show mesophases. From optical, thermodynamic and X-ray studies, I ventured to propose the following structure for the mesophase: the discs are stacked one on top of the other *aperiodically* to form liquid-like columns, the different columns constituting a two-dimensional array<sup>1</sup> (Fig. 5a). In other words, the structure has translational periodicity in two dimensions but not in the third. Other disc-shaped mesogens—the hexa-substituted alkoxy and

alkanoyloxy triphenylenes (Fig. 2(a))—were reported soon afterwards by the Paris and Bordeaux groups<sup>2,3</sup>, and the basic columnar structure was confirmed by the excellent X-ray work of Levelut<sup>4,5</sup>.

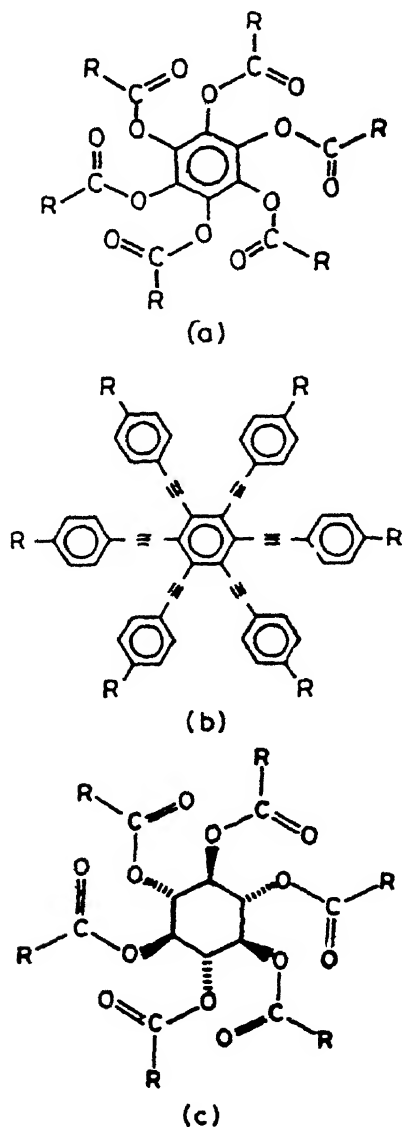


FIG 1(a) Hexa-*n*-alkanoyloxybenzene (after Chandrasekhar *et al*<sup>1</sup>), (b) hexakis (4-octyl-phenyl) ethynyl) benzene (after Kohne and Praefcke<sup>68</sup>), (c) hexa-*n*-alkanoates of scylloinositol (after Kohne and Praefcke<sup>69</sup>)

It is of interest to note that the benzene derivatives were first synthesized as far back as 1937 by Backer and Van der Baan<sup>6</sup> at the University of Groningen, but the authors did not report the mesogenic character of these compounds. However, in 1983 WH de Jeu<sup>7</sup>, who works in Groningen, found to his pleasant surprise that the materials were still in



stock in the Department of Organic Chemistry, and sure enough they did show the expected mesomorphic properties.

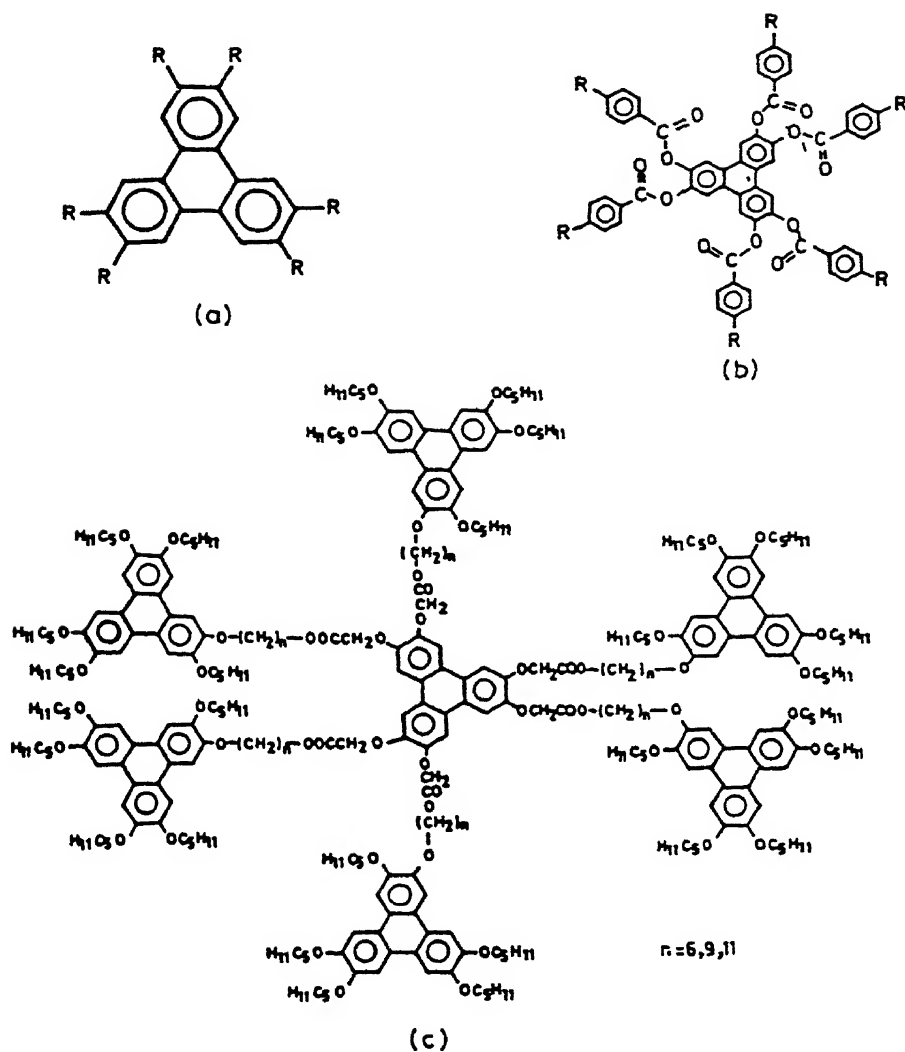


FIG 2 (a) hexa-substituted alkanoyloxy and alkoxy triphenylene (after Billard *et al*<sup>2</sup>; Destrade *et al*<sup>3</sup>) (b) hexa-*n*-alkyl and alkoxybenzoates of triphenylene (after Tinh *et al*<sup>70</sup>; Destrade *et al*<sup>71</sup>); (c) 'star-like' heptameric triphenylene derivatives (after Diele *et al*<sup>20</sup>).

Another early observation was that of Eaborn and Hartshorne<sup>8</sup>, who, in 1955, found that di-isobutyl silane diol (DIBSD) exhibited a mesophase which they were unable to classify as belonging to any of the then known liquid crystal types. Much later (1980), Bunning *et al.*<sup>9</sup> carried out miscibility studies of this compound with one of our benzene derivatives, benzene hexa-*n* heptanoate, and proved conclusively that the mesophase of DIBSD is a columnar liquid crystal. Strictly speaking DIBSD cannot

exactly be described as discotic, but the authors have suggested that it forms a dimer or a polymer which favours the occurrence of a columnar phase.

Vill<sup>10</sup> has pointed out that a sodium salt of diphenyl acetic acid synthesized by Vorlander<sup>11</sup> in 1910 was later shown to exhibit a hexagonal phase<sup>12</sup>. Again, as in the case of DIBSD, the molecule itself is not exactly discotic, and a more detailed study is probably necessary to verify that the mesophase is, in fact columnar.

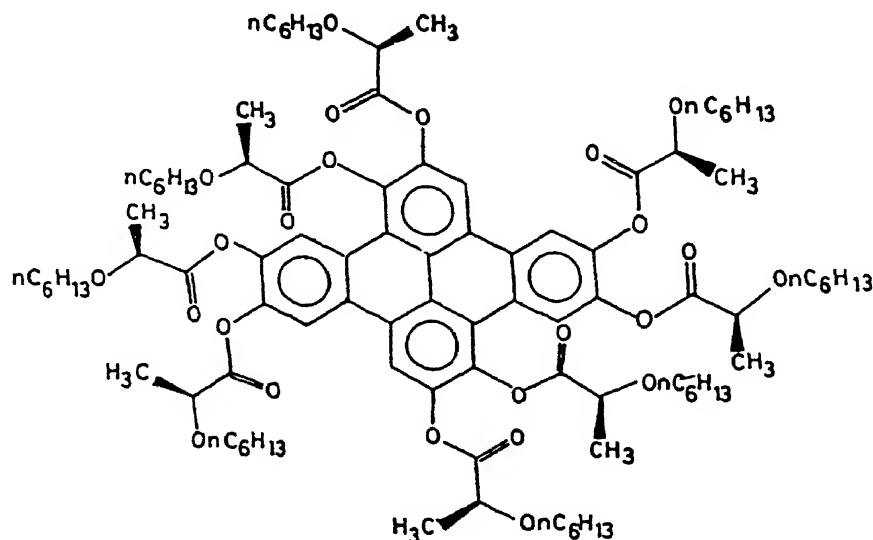


FIG 3 1,2,5,6,9,12, 13-Octa-(S-2-hexyloxypropanoyloxy)-dibenzo(-dibenzo-(e, 1)-pyrene, which forms a ferroelectric columnar liquid crystal (after Bock and Helfrich<sup>21</sup>).

The subject of discotic liquid crystals has grown rapidly during the last decade. Hinov<sup>13</sup> prepared a bibliography of 111 papers on discotics published up to 1984. Very recently, Vill and Thiem<sup>14</sup> have set up a Liquid Crystal Data Bank. Dr Volkmar Vill informs me that approximately 1000 discotic compounds are known to date and some 800 papers have been published on the chemistry and physics of these materials.<sup>15</sup> There is also considerable interest in discotic polymers, a new class of polymer liquid crystals composed of discotic monomer units.

In this lecture we shall review very briefly the major advances in the field, with emphasis on the structural aspects rather than the physics of these systems. Thus, we shall not touch upon the elastic and viscous behaviour, the continuum theory, the theory of defects, molecular

statistical models, the theory of phase transitions in discotics etc.<sup>16-19</sup>, but confine ourselves mainly to a description of the molecules and the mesophase structures. As a further illustration of the role of molecular engineering in this field, we shall also discuss very briefly the steps that led to the important discovery of the biaxial nematic phase in simple thermotropic systems.

## DESCRIPTION OF THE LIQUID CRYSTALLINE STRUCTURES

Generally speaking, discotic mesogens have flat or nearly flat cores surrounded by six or eight (sometimes four) long chain substituents. Some examples are presented in Figs. 1-3. Fig 2(c) gives the structure of an oligometric triphenylene derivative showing a hexagonal columnar

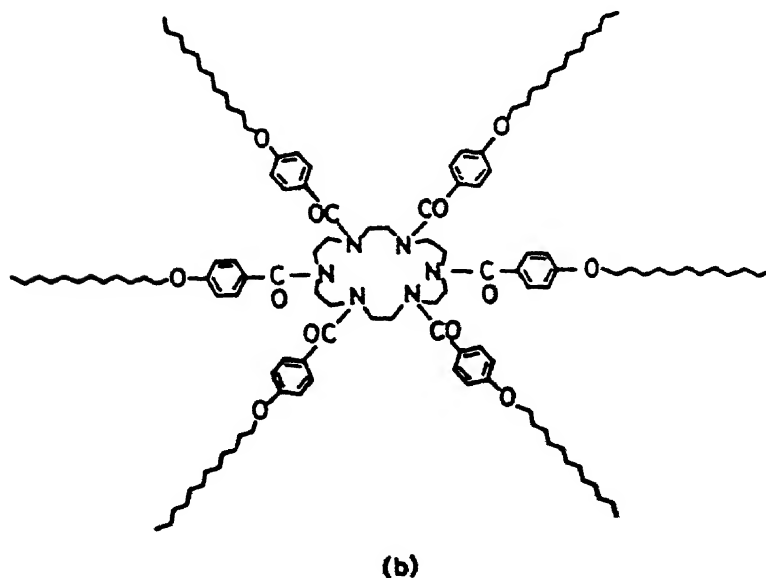
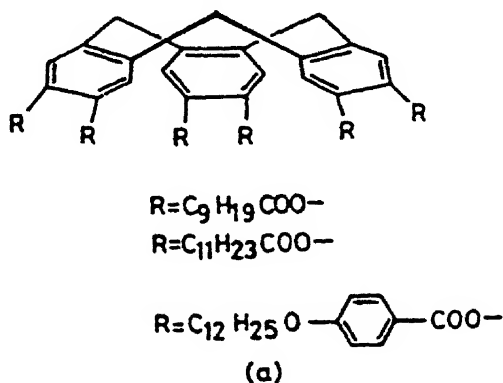


FIG 4 Cyclotricatechylene hexaesters (after Malthete and Collet<sup>22</sup>); (b) hexa-(*p*-*n*-dodecyloxybenzoyl) derivatives of macrocyclic polyamines (after Lehn *et al.*<sup>24</sup>).

phase<sup>20</sup>. The molecule of Fig. 3 is noteworthy in that it is the first and, to date, only example of a compound exhibiting a ferroelectric columnar phase<sup>21</sup>. Columnar phases are also formed when the flat core is replaced by a conical or pyramidal shaped one<sup>22,23</sup> [(Fig. 4(a)], and in rare cases, even when the central core is absent altogether, as in certain macrocyclic molecules<sup>24</sup> [Fig 4(b)].

Several variants of the columnar structure have been identified (Fig. 5a). High resolution synchrotron X-ray studies on a few compounds have thrown light on the details of these structures. The studies were carried out on very well oriented monodomain samples obtained by preparing freely suspended liquid crystal strands, typically about 200  $\mu\text{m}$  in diameter and 1.5-2 mm long, with the column axis parallel to the axis of the strand<sup>25-29</sup>. The principal results of these investigations are summarized below.

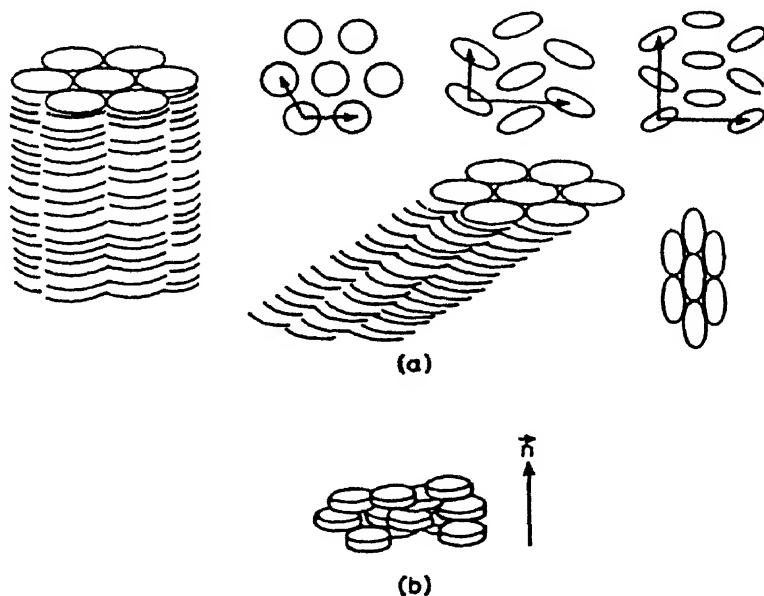


FIG 5 (a) Plan views of the two-dimensional lattices of the columnar phases of discotic molecules. Ellipses denote discs that are tilted with respect to the basal plane (after Lavelut<sup>5</sup>); (b) the discotic nematic phase.

- (i) The correlation length of the two-dimensional lattice of the columnar phase of  $(\text{C}_{13} \text{H}_{27}\text{COO})_6$ -truxene is greater than 4000  $\text{\AA}$ , the lower limit being set by the instrumental resolution<sup>29</sup>. On the other hand, within each column the flat molecular cores form an orientationally ordered one-dimensional liquid, while the hydrocarbon chain surrounding the core are in a highly disordered state. The marked difference between the ordering of the cores and the chains is also confirmed by deuterium NMR spectroscopy<sup>30</sup>.

The structure of this columnar phase is the same as that proposed originally by us for the mesophase of the benzene compounds.<sup>1</sup> The possibility of such a structure was, in fact, envisaged by Landau while discussing the stability of 'low-dimensional solids'. Specifically Landau was considering a three-dimensional body in which the density is a periodic function of two dimensions only, and concluded as follows<sup>31</sup>: "*Thus, bodies having such a structure could in theory exist, but it is not known whether they do in fact exist in Nature*". The work on discotics has provided the first direct experimental verification of Landau's prediction.

- (ii)  $(C_{12}H_{25}COO)_6$ -triphenylene exhibits a transition from a hexagonal ( $D_h$ ) to a rectangular ( $D_r$ ) columnar phase. The transition, which is weakly first-order, is associated with a small distortion of the lattice, consistent with a herringbone arrangement in the rectangular structure with only the core of the molecule tilted with respect to the

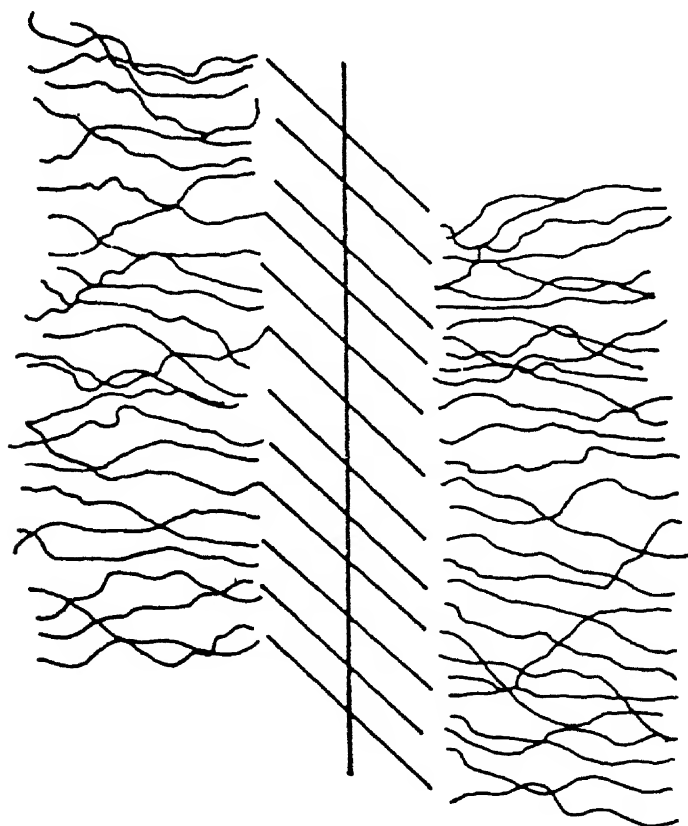


FIG 6 Schematic representation of the structure of a column in which the cores are tilted with respect to the column axis and the chains are in a highly disordered state (after Frank and Chandrasekhar<sup>32</sup>).

column axis, in accordance with a model proposed by F C Frank and myself<sup>32</sup> (Fig. 6). However, high resolution synchrotron X-ray studies on monodomain discotic strands have established that the tilt of the molecular core persists in the  $D_h$  phase as well, except that the tilt in neighbouring columns are rotationally uncorrelated, i.e. they are free to assume different azimuthal angles<sup>26,27</sup>. Interestingly, therefore, the  $D_h$ - $D_r$  transition may be looked upon as an orientational order-disorder transition.

- (iii) Hexa-hexyl-thiotriphenylene shows a transition from a hexagonal 'ordered' phase to a hexagonal 'disordered' one. The former is a phase in which there is regularity in the stacking of the triphenylene cores in each column, and the latter one in which the column is liquid-like. X-ray studies reveal that in the ordered phase there is a helicoidal stacking of the cores within each column, the helical spacing being incommensurate with the intermolecular spacing<sup>28</sup>. In addition, a three-column superlattice develops as a result of the frustration caused by molecular interdigitation in triangular symmetry. Ideally, if there is no intercolumn interaction, true long-range order cannot exist within a column because of the Peierls-Landau instability. The existence of a regular periodicity in the stacking in each column therefore implies that neighbouring columns must be in register. Thus ordered columnar phases can probably be compared with the highly ordered smectic phases of rod-like molecules e.g., smectics B, E, G, H etc. which possess three-dimensional positional order. However, further high resolution studies are necessary before general conclusions can be drawn.

A nematic type of phase ( $N_D$ ) has also been identified (Fig. 5b). The director now denotes the preferred axis of orientation of the disc-normals (or short molecular axes). Hence, the medium is optically negative, unlike the classical nematic which is optically positive. A twisted nematic (or cholesteric) phase with the helical axis normal to the director has also been observed<sup>33</sup>.

A smectic-like lamellar discotic phase has been reported<sup>34-37</sup>. The suggestion has been made<sup>35,36</sup> that the phase has a tilted smectic C type of structure, but the disposition of the molecules in each layer does not appear to have been resolved conclusively.

## DISCOTIC METALLOMESOGENS

Mesogens containing metal atoms are now attracting a great deal of attention mainly because of their unusual electrical and magnetic properties. The first discotic metallomesogens were reported by Giroud-Godquin and Billard<sup>34</sup>, and since then a variety of metal complexes exhibiting discotic phases have been synthesized<sup>38</sup>. Two examples are given in Fig. 7.

Available electrical conductivity data<sup>39,40</sup> on metallomesogens indicate that the columnar phases are molecular semiconductors. Calculations show that the hopping probability for the electron is very much greater along the column axis than transverse to it, as is to be expected.

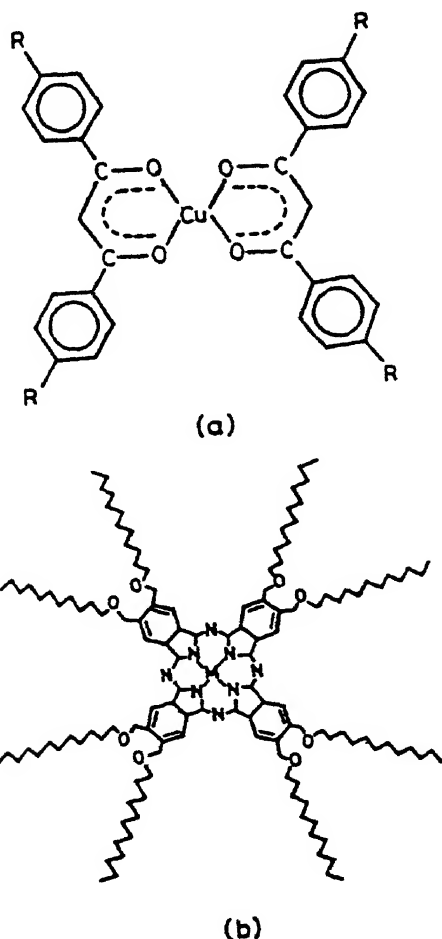


FIG 7 Metallomesogens (a) bis(*p-n*-decylbenzoyl) methanato copper (II) (Giroud-Godquin and Billard<sup>34</sup>); and (b) octa-substituted metallophthalocyanine (after Piechoki *et al.*<sup>72</sup>).

Eastman *et al.*<sup>41</sup> have carried out ESR studies on single crystals of the copper complex of Fig. 7(a). From an analysis of the data, the authors have concluded that the spectra have the features associated with a spin  $1/2$  one-dimensional Heisenberg antiferromagnet, and that the exchange interactions are quite significant (i.e., an appreciable degree of antiferromagnetic long range order persists) even in the discotic phase.

### THE BIAxIAL NEMATIC LIQUID CRYSTAL

The biaxial nematic ( $N_b$ ) phase was first identified by Yu and Saupe<sup>42</sup> in a ternary amphiphilic system composed of potassium laurate, 1-decanol and  $D_2O$ . In such systems, the constituent units are molecular aggregates, called micelles, whose size, shape and number density are sensitive functions of concentration and temperature: the  $N_b$  phase was found to occur over a range concentration/temperature. There are obvious advantages in having a single-component, low molecular weight thermotropic  $N_b$  phase. Some years ago<sup>43</sup>, I put forward the suggestion that a convenient method of achieving this would be by "*bridging the gap between rod-like and disc-like mesogens*", i.e., by preparing a molecule that combines the features of the rod and the disc. This idea has proved to be efficacious, and the  $N_b$  phase has since been observed in relatively simple compounds<sup>44</sup>. Two examples are presented in Fig. 8. The copper complex of Fig. 8(a) can be seen to be an elongated version of the discotic complex of Fig. 7(a). It was also the first paramagnetic nematogen to be studied. Similarly, the compound of Fig. 8(b) evidently combines the features of the rod and the disc.

We shall discuss the results for the copper complex of Fig 8(a) in some detail. The biaxiality of the mesophase was inferred from the following observations<sup>46-48,73</sup>

- (i) The appearance of the usual nematic schlieren texture, when viewed under the polarizing microscope, except that often the pattern consisted almost entirely of disclinations of strength  $|s| = 1/2$  (two-brush disclinations). This may be taken to be evidence of biaxiality. Since the escape mechanism for the  $|s|=1$  disclinations does not, in principle, eliminate the singularity in a biaxial nematic, as it does in uniaxial nematic.<sup>49,50</sup>, the four-brush disclinations probably become energetically unfavourable.



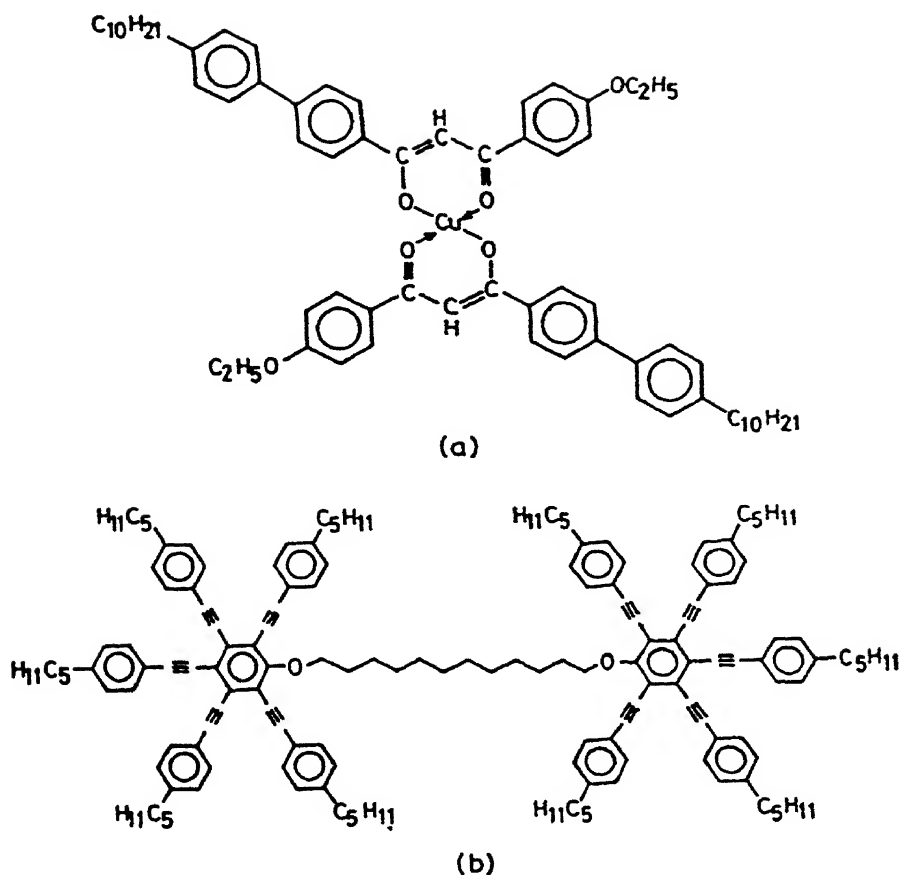


FIG 8 Molecules which exhibit the biaxial nematic phase (a) bis-1(*p*-*n*-decylbiphenyl)-3-(*p*-ethoxy phenyl) propane-1, 3-dionato copper (II) (Chandrasekhar *et al* <sup>45</sup>), and (b) 1,12-bis(pentakis(4-pentyl-phenyl)ethynyl)phenyloxy)dodecane (after Praefcke *et al.*<sup>52</sup>)

- (ii) The occasional (and unpredictable) appearance of zig-zag disclinations<sup>51</sup>, though this could not be taken to be conclusive proof of biaxiality.
- (iii) Conoscopic observations using thick films (~125  $\mu\text{m}$ ). Homeotropically aligned sample were obtained by the combined effect of silane coating on the internal surfaces of the cover slips, and a 3 kHz AC electric field applied to electrodes coated on the external surfaces. The alignment was checked by measuring the intensity of light transmitted by the sample between crossed polaroids, under orthoscopic conditions using a He-Ne laser and a photodiode. For perfect alignment there was almost complete

extinction, and the transmitted intensity was equal to that for the isotropic phase. No electrohydrodynamic motion was seen in pure samples. There was evidence of some chemical decomposition on repeated heating of the material and therefore only fresh samples were used for the experiments. All the conoscopic observations were found to be reproducible with well aligned samples in freshly prepared cells. The conoscopic pattern was independent of the applied voltage for voltages greater than the saturation value. We were able to demonstrate (a) the biaxiality of the nematic phase of the pure complex, (b) a reversible uniaxial-biaxial ( $N_a$ - $N_b$ ) transition in a binary mixture of the complex with the uniaxial nematogen 4-n-pentyl-4-cyano-p-terphenyl, and (c) the small variation of the biaxiality with temperature near this transition. We also obtained the  $I$ - $N_a$ - $N_b$  phase diagram for this binary system.

- (iv) X-ray diffraction photographs of samples aligned in a magnetic field of 1.8T. Intensity scans using a microdensitometer revealed an additional pair of diffuse peaks in the equatorial scans, as would be expected of an orthorhombic fluid. Strictly speaking, this in itself cannot be considered to be conclusive proof of long-range biaxial nematic order; a complete X-ray study would require measurements on monodomain specimens obtained by applying crossed electric and magnetic fields. However, the present X ray evidence, along with the optical observations described earlier characterizes the phase to be a biaxial nematic.

Similar optical and X-ray studies were used by Praefcke *et al.*<sup>52</sup> to identify the  $N_b$  phase of the compound shown in Fig. 8(b).

A number of important theoretical studies concerning the  $N_b$  phase have been discussed—statistical theories, continuum theories, topological theories of defects. etc.(for a full list of references see ref. 46). For example, the orthorhombic  $N_b$  phase has 15 elastic constant and 15 viscosity coefficients<sup>53-56</sup> (as compared with 3 elastic constants and 5 viscosity coefficients for the uniaxial case. Again, a remarkable conclusion of the homotopy theory is that the usual law of coalescence of defects breaks down in the  $N_b$  phase, the combination rule now being non-Abelian. Moreover, there can arise an entanglement of disclinations which should give rise to what Toulouse calls ‘topological rigidity’<sup>49,57-60</sup>.

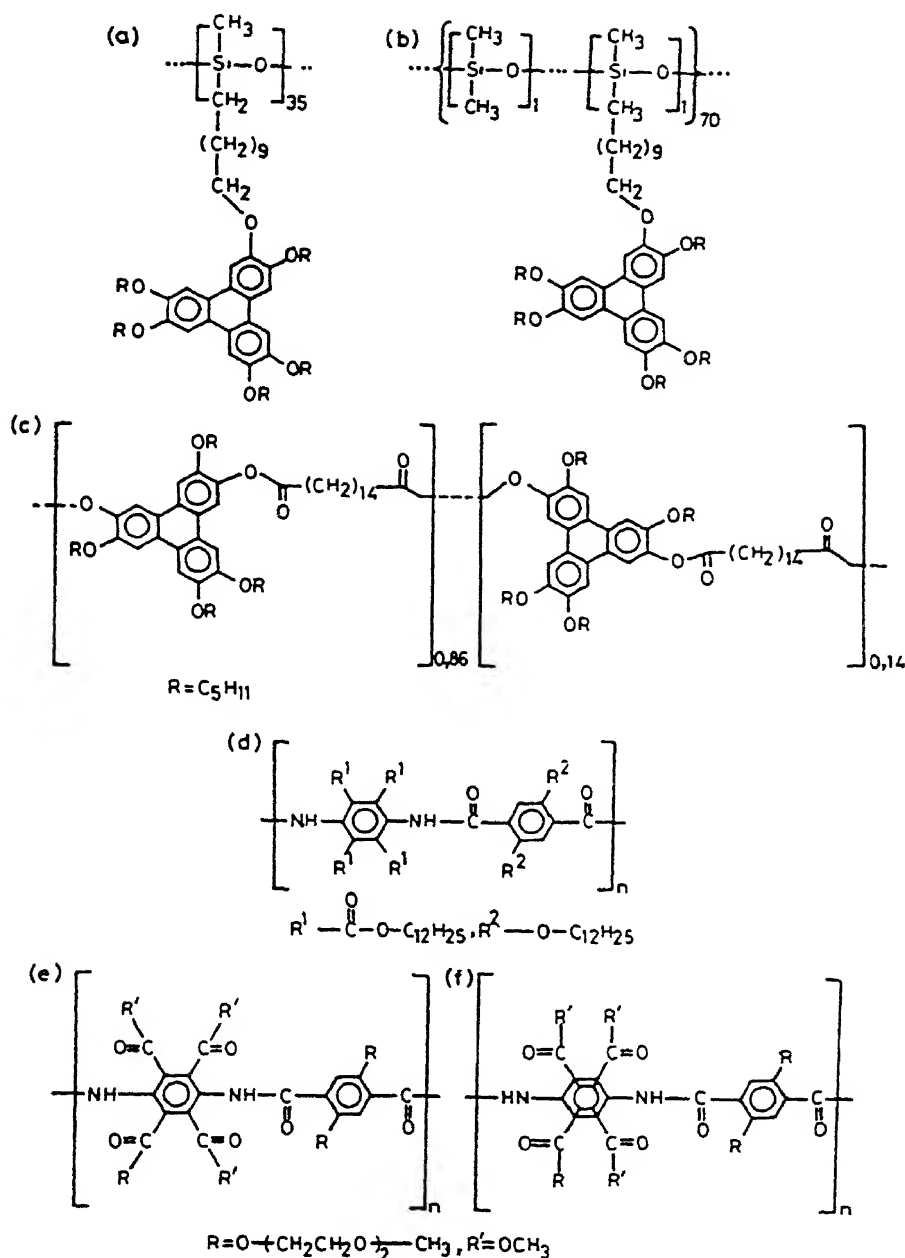


FIG 9 Discotic polymers: transition temperatures in °C (a) annealed sample, G-19-CD 39°I (after Kreuder and Ringsdorf<sup>61</sup>); (b) annealed sample, G-29° D 36° I (After Kreuder and Ringsdorf<sup>61</sup>), (c) G 60° D 150°I (Herrmann-Schönherr *et al*<sup>62</sup>); (d) D<sub>1</sub> 67° D<sub>2</sub> 130°I (after Herrmann-Schönherr *et al*<sup>63</sup>); (e) K 179°N 262°I (after Ebert *et al*<sup>64</sup>); and (f) K 131° N<sub>σ</sub> 216°I (Ebert *et al*<sup>64</sup>). G= glass, K-crystal. D- discotic columnar. N<sub>σ</sub> - sanidic nematic, I = isotropic.

These and other ideas have yet to be investigated experimentally. The availability of a simple thermotropic biaxial nematic should make it conveniently possible to test some of these predictions.

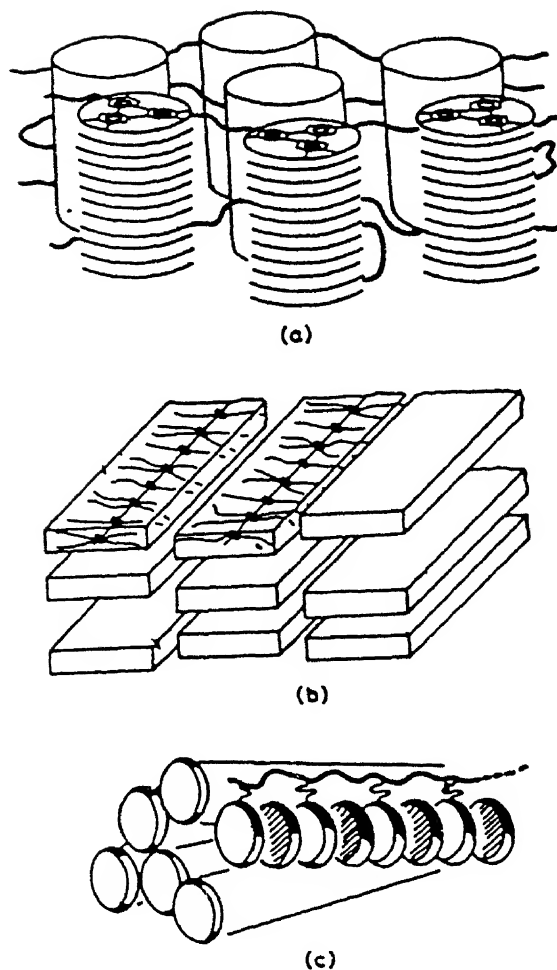


Fig 10 Mesophases of discotic polymers: (a) the hexagonal columnar phase (Hermann-Schönherr *et al*<sup>62</sup> Huser *et al*<sup>66</sup>); (b) the smectic nematic phase ( $N_\sigma$ ) (Ebert *et al*<sup>64</sup>), (c) the columnar nematic phase (Ringsdorf and Wustefeld<sup>67</sup>).

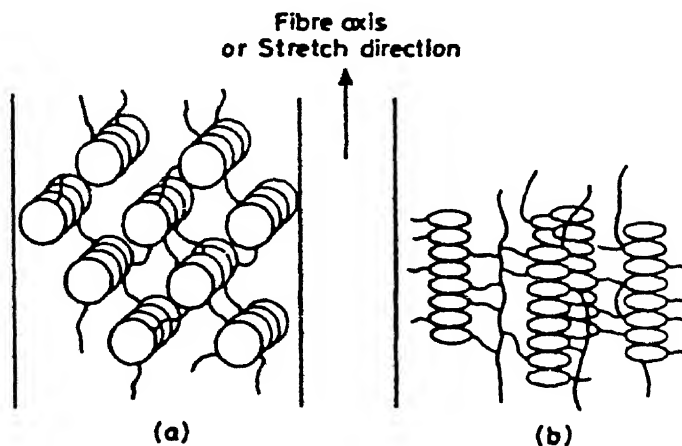


FIG 11 Orientation of the molecules in stretched samples (a) main chain discotic polymers: and (b) side group discotic polymers (Huser *et al*<sup>66</sup>).

## DISCOTIC POLYMER LIQUID CRYSTALS

Finally we turn our attention briefly to a new class of liquid crystal polymers, viz., discotic polymers<sup>61-67</sup>. The basic monomer units are disc-shaped mesogens which are attached to the polymer backbone in the main chain itself or as side groups (Fig. 9). The types of mesophases that have so far been identified are illustrated in Fig. 10. A polyester with triphenylene as the repeating unit in the main chain separated by flexible spacers<sup>62</sup>, or with teriphenylene units attached as side groups to the polymer chain *via* flexible spacers<sup>66</sup> forms a hexagonal columnar mesophase [(Fig. 10(a))]. On the other hand, rigid aromatic polyamides and polyesters with disc-shaped units in the main chain form a smectic (or board-like) nematic<sup>64</sup>, with the boards stacked parallel to one another (Fig. 10(b)). The addition of electron acceptor molecules to discotic polymers result in the formation of charge transfer complexes which stabilize (or in certain, non-mesomorphic materials, induce) mesophases<sup>67</sup>. A new type of induced mesophase in such a system is the 'columnar nematic' [(Fig 10(c))].

X-ray diffraction studies of the macroscopic alignment of the columnar structures produced by stretching the film, or by drawing fibres (about 1 m long and several microns thick) from the melt, have been reported<sup>62,66</sup>. Hüser, Pakula and Spiess<sup>66</sup> have demonstrated that almost perfect alignment can be achieved by proper mechanical and thermal treatment. In the main chain polymer, the columns are oriented perpendicular to the stretch direction (or the chain direction) whereas in the side group polymer they are oriented parallel to it (Fig.11). A detailed study of the mechanical and other physical properties of these oriented polymers is evidently of great interest.

## ACKNOWLEDGEMENT

I am very grateful to Professor H Ringsdorf, Dr P Schuhmacher, Dr R Wustefeld, Dr I G Voigt-Martin, Professor K Praefcke, Professor J H Wendorff, Professor H Spiess, Dr Christine Boeffel, Professor G Heppke and their colleagues for bringing me up to date, particularly in the field of discotic polymers, and to Dr Volkmar Vill for many useful discussions on the LIQCRYST Daten bank.

## References

1. S Chandrasekhar, B K Sadashiva and K A Suresh *Pramana* 7 (1977) 471

2. J Bilard, O C Dubois, N H Tinh and A Zahnn *Nouv J Chim* **2** (1978) 535
3. C Destrade, M C Mondon and J Malthete *J Phys (Paris)* **40** (1979) C3-17
4. A M Levelut *J Phys Lett(Paris)* **40** (1979) L81
5. A M Levelut *J Chim Phys* **88** (1983) 149
6. H J Backer and S Van der Baan *Rec Trav Chim* **56** (1937) 1161
7. W H de Jeu *Phil Trans R Soc A* **309** (1983) 102
8. C Eaborn and N H Hartshorne *J Chem Soc* (1955) 549
9. J C Bunning, J W Gooby, G W Gray and J B Lydon *Proc Conf Liq Cryst One and Two-Dimensional order, Garmisch Partenkirchen* (Eds W Helfrich and G Heppke) Springer-Verlag (1980) p. 397
10. V Vill *Condensed Matter News* **1** (5) (1992) 25
11. D Vorlander *Ber Dtsch Chem Ges* **43** (1910) 3120
12. D Demus H Sackmann and K Siebert *Wiss S Univ Math Nat R* **19**(5) (1970) 47
13. H P Hinov *Mol Cryst Liq Cryst* **136** (1986) 221
14. V Vill and J Thiem *Liq Cryst Datenbank Flüssigkristalliner Verbindungen* Springer-Verlag (1992)
15. V Vill (*Personal Commun*)
16. S Chandrasekhar *Advances in Liquid Crystals* **5** (Ed G H Brown) Academic (1982) p 247
17. S Chandrasekhar *Phil Trans R Soc A* **302** (1983) 93
18. S Chandrasekhar and G S Ranganath *Rep Prog Phys* **53** (1990) 57
19. S Chadrasekhar Plenary Lecture at the 14th Int Liq Cryst Conf Pisa June 1992 (*to be published*)
20. S Diele, A Liebmam, U Nutz, H Ringsdorf and P Schuhmacher *14th Int Liq Cryst Conf Pisa* June 1992 Abstr. No. A-P 83 (1992)
21. H Bock and W Helfrich *14th Int Liq Cryst Conf Pisa* (1992) Abstr No. CP. 99
22. J Malthete and A Collet *Nouv J Chim* **9** (1985) 151
23. H Zimmerman, Poupko, Z Luz and J Billard *Z Nature A* **40** (1985) 149
24. J M Lehn, J Malthete and A M Levelut *J Chem Soc Chem Commun* (1985) 1764
25. D H Van Winkle and N A Clark *Phys Rev Lett* **48** (1982) 1407
26. C R Safinya, N A Clark, K S Liang, W A Varady and L C Chiang *Mol Cryst Liq Cryst* **123** (1985) 205
27. C R Safinya, K S Liang, W A Varady, N A Clark and G Anderson *Phys Rev Lett* **53** (1984) 1172

28. E Fontes, P A Heiney and W H de Jeu *Phys Rev Lett* **61** (1988) 1202
29. E Fontes, P A Heiney, M Ohba, J N Haseltine and B A Smith *Phys Rev A* **37** (1988) 1329
30. D Goldfrab, Z Luz and H Zimmermann *J Phys Paris* **42** (1981) 1303
31. L D Landau and E M Lifshitz *Statistical Physics part 11* (3rd Ed) Pergamon (1980) p. 435
32. F C Frank and S Chandrasekhar *J Phys Paris* **41** (1980) 1285
33. G Heppke, D Kruerke and B Sabaschus *14th Int Liq Cryst Conf Pisa Abstr* B-P 17, C Destrade N H Tinh J Malthete and J Jacques *Phys Lett A* **79** (1980) 189
34. A M Giroud-Godquin and J Billard *Mol Cryst Liq Cryst* **66** (1981) 147
35. K Ohta, H Muroki, A Takagi, K I Hatada, H Ema, Yamamoto and K Matsuzuki *Mol Cryst Liq Cryst* **140** (1986) 131
36. H Sakashita, A Nishitani, Y Sumiya, H Terauchi, K Ohta and K Matsuzuki *Mol Cryst Liq Cryst* **163** (1988) 211
37. A C Ribeiro, A F Martins and A M Giroud-Godquin *Mol Cryst Liq Cryst Lett* **5** (1988) 133
38. A M Giroud-Godquin and P Maitlis *Angew Chem Int (Edn Engl)* **30** (1991) 375
39. J F Van der Pol, E Neeleman, J W Zwikker, R J M Nolte W Drnith Aerts R Visser and S J Picken *Liq Cryst* **6** (1989) 577
40. Z Belarbi, M Maitrot, K Ohta, J Simon, J J Andre and P Petit *Chem Phys Lett* **143** (1988) 400; M J Cook, M F Daniel, K J Harrison, N B Mckeown and A J Thomson *J chem Soc chem Commun*(1987) p 1086
41. M P Eastman, M L Horng, B Freiha and K W Sheu *Liq Cryst* **2** (1987) 223
42. L J Yu and A Saupe *Phys Rev Lett* **45** (1980) 1000
43. S Chandrasekhar *Mol Cryst Liq Cryst* **124** (1985) 1
44. K Praefcke, B Kohne, B Gundogan, D Singer, D Demus, S Diele, G Pelzi and U Bakowsky *Mol Cryst Liq Cryst* **198** (1991) 393
45. S Chandrasekhar, B K Sadashiva S Ramesh and B S Srikanta *Pramana J Phys* **27**(1986) L713
46. S Chandrasekhar, B R Ratna, B K Sadashiva and V N Raja *Mol Cryst Liq Cryst* **165** (1990) 123
47. S Chandrasekhar, V N Raja, B K Sadashiva *Mol Cryst Liq Cryst Lett* **7** (1990) 65
48. S Chandrasekhar *Sir Charles Frank—An Eightieth Birthday Tribute* (1991) p. 254
49. N D Mermin *Rev Mod Phys* **51** (1979) 591
50. S Chandrasekhar and G S Ranganath *Adv Phys* **35** (1986) 507

51. Y Galerne and L Liebert *Phys Rev Lett* **55** (1985) 2449
52. K Praefcke, B Kohne, B Gundogan, D Demus, S Diele and G Pelzl *Mol Cryst Liq Cryst Lett* **7** (1990) 27
53. A Saupe *J Chem Phys* **75** (1981) 5118
54. U D Kini *Mol Cryst Liq Cryst* **108** (1984) 71
55. U D Kini and S Chandrasekhar *Physica* **156A** (1989) 364
56. U D Kini and S Chandrasekhar *Mol Cryst Liq Cryst* **179** (1990) 27
57. C Toulouse *J Phys Lett (Paris)* **38** (1990) L67
58. V Poenaru and G Toulouse *J Phys (Paris)* **8** (1977) 887
59. H R Trebin *Adv Phys* **31** (1982) 195
60. M Kleman *Points, Lines and Walls* Wiley New York (1983)
61. W Kreuder and H Ringsdorf *Makrom Chem Rapid Commun* **4** (1983) 807
62. O Herrmann-Schönherr, J H Wendorff, W Kreuder and H Ringsdorf *Makrom Chem Rapid Commun* (1986)
63. O Herrmann-Schönherr, J H Wendorff, H Ringsdorf and P Tschirner *Makrom Chem Rapid Commun* **7** (1986) 791
64. M Ebert, O Herrmann-Schönherr, J H Wendorff, H Ringsdorf and P Tschirner *Makrom Chem Rapid Commun* **9** (1988) 445
65. B Kohne, K Praefcke, H Ringsdorf and P Tschirner *Liq Cryst* **4** (1989) 165
66. B Hüser, T Pakula, H W Spiess *Macromolecules* **22** (1960)
67. H Ringsdorf and R Wustefeld *Mol Chem for Electronics* The Royal Society (1990) p 23)
68. B Kohne and K Praefcke *Chimia* **41** (1987) 196
69. B Kohne and K Praefcke *Angew Chem* **96** (1987) 70
70. N H Tinh, C Destrade and H Gasparoux *Phys Lett* **72A** (1979) 251
71. C Destrade, M C Bernaud, H Gasparoux, A M Levelut and N H Tinh *Proc Int Conf Liq Cryst (Bangalore)* (1980) Ed. Chandrasekhar S. Hyden p. 29
72. C Piechoki, J Simon, A Skoulios, D Guillon and P Weber *J Am Chem Soc* **104** (1982) 5245
73. S Chandrasekhar *Plenary Lecture* Solid State Physics Conference Institute of Physics, Bristol, U K 16-18 December 1987





Pratap Narain Srivastava (b. 1 July 1927) did D.Phil (1951) from University of Allahabad. He is Member of Planning Commission, Government of India and Vice-Chancellor of Jawaharlal Nehru University (JNU), New Delhi

Srivastava has done much work on the radiation effects on adult mice during pre-natal and post-natal development. For a long time, the relative biological effectiveness of tritium with respect to  $^{60}\text{Co}$  had been considered to be 1. His work, along with others, has led to its upward revision. His work on the mechanism of action of radioprotective drugs has

received recognition. His recent work on entrapping radioprotective drugs with liposomes and then targeting them to particular tissues by linking specific antibodies on to the surface of liposome has great potential for reducing the quantity of toxic radioprotective drugs.

Srivastava is Fellow (President, Biological Sciences, 1978) of National Academy of Sciences (India), Royal Microscopical Society (UK), and Linnean Society (London) has been Treasurer of INSA (1985-88). He is the recipient of UGC National Lecturer (1977-78), UGC National Fellow (1979-81), Professor Emeritus for Life, JNU; FICCI Award (1984), Gold Medal, Institute of Oriental Philosophy, Japan (1986), Aryabhata Medal (INSA) (1989), K N Bahl Memorial Gold Medal, (Indian Society of Biosciences) (1990); Atma Ram Award (Ministry of Human Resource Development) (1991); Professor Distinguido (La Universidad Externado De Colombia, Bogota) (1992). He is presently President of Indian Science Congress (1993-94).

*Pratap Narayan was elected to the fellowship of the Academy in 1977.*

**Address:** Nuclear Science Centre, Jawaharlal Nehru University Campus, New Delhi 110 067, Tel. (011) 6893955, 6892601, 6892603, 6893467, 6891410, Telex: 031- 73275 NSC IN; Fax: 91-11-6893666; E-mail uunet! shakti! vikram!nsc; Res: F1/128, Asiad Games Village, Madanlal Block, New Delhi-110 049, Tel: (011) 6493322.

## **CANCER AND RADIATION THERAPY: PRESENT STATUS**

P N SRIVASTAVA

In a very simple language cancer can be described as a state when the normal growth and differentiation go awry resulting into malignant tumours. Cancerous cells continue to divide; live longer than normal cells; appear not to undergo the programmed cell death characteristic of normal differentiation to non-dividing, terminating differentiated cells. Cancers may be caused by many factors such as chemicals, radiation and hormones. Chemicals are generally associated with skin, lung and bladder tumours, radiation with haematological malignancies, leukaemia, lymphoma as well as thyroid carcinoma, while hormones are associated with cancers in breast and uterus. Growth of malignant cancer depends not just on factors causing cancer but also on body's response to it as well.

It was in the fitness of things that 1989 Nobel Prize in Medicine was awarded to J Michael Bishop and Harold Varmus of the University of California for their work on oncogene hypothesis in 1976. Their work suggested that normal genes (referred to as proto-oncogenes) controlling growth, development and differentiation somehow become misdirected and mutate into oncogenes in neoplastic cancer cells. While remembering the Nobel Prize winners, let us not forget Alexander Haddow who had indicated as far back as in 1937 that cancer growth arises as a somatic mutation. Oncogenes have been referred to acting like red and green traffic lights regulating cell growth. When they are disrupted, it is like green light getting stuck allowing a tumour to grow unchecked. The discovery has widened our knowledge and insight into the complicated signal systems which govern the normal growth of cells. The importance of Bishop and Varmus' work lies in the fact that we have now a completely different view on how cancer can originate. It is still not very well known how a normal gene gets triggered into becoming an oncogene. Once we know this process, we may also be able to stop this triggering mechanism hopefully some day. We also know by experience that when treating a tumour we often sacrifice and alter many normal cells which cannot be helped. The destruction of the normal cells is the reason for the significant side effects associated with current cancer therapies.

## CANCER AND AGE

Cancer is a disease of old age except some such as leukaemia and bone cancer. Out of fifty million deaths per year in the world, more than five million are attributed to cancer. It is estimated that in another ten to fifteen years, cancer deaths in the United States of America alone may account for eight million. The main reasons for this are changes in the health spectrum and demographic structure, changes in life style and changes in environment. More than half the world's population now lives in countries where cancer is among the top ranking causes of death. In Japan, it is number one and in the United States of America, it is number two. This is because longevity is now highest in Japan. It may also be worthwhile to note the age distribution of population between the age of ten to forty years. It is 40% in Africa, 31% in Asia and 12% in Europe and North-America. However, the average age of cancer patients in Europe and North-America is 55.7 years, while in Asia and Africa it is 45.9 and 35.9 years respectively. This must be because the standard of health is poorer in Asia and Africa.

## INCIDENCE OF CANCER AND DEATHS

In the United States of America death due to cancer was 107/100,000 population in 1930 and it increased to 180/100,000 in 1980. The increase must also be related to the age since in the fifty years period longevity has also increased. However, when the figure is corrected for longevity factor, even then it comes to 160/100,000 population which also is a significant increase. Unfortunately, a correct estimate of this is not available in India. It is estimated that the incidence of cancer was 80/100,000 in 1947 and has increased to 120/100,000 population at present and is going to increase two to three folds by the year 2000 A.D. Longevity has increased in India also, from 37 years in 1947 to 58 years now.

## RADIATION THERAPY

For a long time radiobiologists received very little attention from radiotherapists but the situation changed during the last decade and a half. As a result, the advances in the field of radiation biology are being received with greater interest and understanding by radiotherapy community. This development has brought in dramatic changes and thus the advances have been faster.

Any tumour can be roughly divided to possess three types of cells, (i) necrotic cells which lie in the centre and are far removed from blood capillaries, (ii) hypoxic cells in the middle, and (iii) oxygenated cells which are at the periphery and are nearest to the blood capillaries. It is also very well established that oxygenated cells are more sensitive to radiation than hypoxic cells. Approximately, thirty to fifty per cent of all cancers need radiation therapy. Single large dose of radiation is harmful, hence fractionated doses are given which is less harmful. Further, in the first dose of radiation the sensitive oxygenated cells are killed when the middle larger hypoxic cells may get converted into oxygenated cells which may be killed by the second dose.

### RADIOPROTECTORS

It has already been stated in the beginning that the significant side effects associated with cancer therapy occur because of the destruction of normal cells which is difficult to avoid. A great deal of effort has been put in by radiobiologists to search a modifier which would be of help to radiotherapists. It is necessary that such a modifier should either selectively potentiate radiation effects on cancer cells without affecting the normal cells or else they should protect the normal cells without protecting the cancer cells.

Many chemicals such as epinephrine, histamine and serotonin are known to act as radioprotectors but they are not true radio-protectors as such since they basically cause vasoconstriction and thus reduce basal metabolic rate. True radioprotectors are sulphhydryl compounds such as cysteine, cysteamine, cystamine, WR-2721, which basically act as free radical scavengers. The difficulty with these drugs have been that to bring about a Dose Reduction Factor (DRF) of 2.0 to 3.0, one has to use 50 to 75 percent of the toxic dose (Sugahara and Srivastava, 1976). Nonetheless, the drugs have been useful. American astronauts perhaps carry cysteamine to be taken in case of emergency.

### 2-MERCAPTOPROPIONYLGLYCINE (MPG)

Our group has done considerable work on MPG since it is non-toxic and gives a DRF of 1.4 at 1% of the toxic dose. This drug was used in Japan for a long time as a detoxicating agent under the trade name Thiola. Sulphydryl compounds have been classified into three classes according to their toxicity and radioprotective action. Most effective protection was obtained by cysteamine and cysteine followed by AET and MPG-amide in

that order. Toxicity of these compounds were generally observed in the range of 0.1 - 2.0 mM (Hikita *et al.*, 1975). On the other hand MPG was found to be non-toxic at 0.02 mM and 15 mM when it gave significant protection. Extensive work on radioprotection by MPG against gamma radiation-induced damage to liver (Saini *et al.*, 1977), bone marrow lymphocytes (Saini *et al.*, 1988), testis (Uma Devi and Saharan, 1978; Saharan and Uma Devi, 1977), small intestine and jejunum (Uma Devi *et al.*, 1978; Uma Devi, 1977; Saharan *et al.*, 1978), thymus (Saini and Uma Devi, 1979a, 1979b; Uma Devi and Saini, 1977), and growth inhibiting effects in utero irradiation in mice (Dev *et al.*, 1982) has been done by our associates. In all these cases an enhanced recovery and accelerated restoration of normal cellular structures and cell counts were observed in case of MPG-pretreated animals as compared to controls.

In spite of the fact, that sulphhydryl compounds are very good radioprotectors, their application in radiotherapy is very much limited because of non-availability on such compounds with twin properties of low toxicity and high DRF. However, two compounds MPG and WR-2721, have been put to clinical tests. While MPG has already undergone clinical trials in Japan (Sugahara and Srivastava, 1976), the clinical trials of WR-2721 are currently in progress (Symposium on Perspective in Radioprotection, March 13-14, 1987, National Bureau of Standards, Bethesda, Maryland, USA). Ayene and Srivastava (1985) have demonstrated radio-sensitization by MPG/Fe complex in microsomes. Enhancement of lipid peroxidation was observed at 0.1 mg/ml of MPG instead of protection. However, MPG in the presence of EDTA gave significantly more protection at all doses of radiation. The complex formation was further confirmed by the exogenous supply of ferrous sulphate during irradiation of microsomes. The addition of Fe ions enhanced the formation of lipid peroxides. A further increase in lipid peroxides was observed in the presence of MPG with an abrupt increase to a dose of 133.2 Gy and a slight increase at higher doses of radiation. Similarly greater protective effect of MPG was observed in erythrocytes in absence of MPG/Fe complex formation (Ayene, Kale and Srivastava, 1988). Protection to erythrocytes at high concentration was also rendered by MPG at both concentrations used. It also exhibited a similar effect in erythrocytes of low concentration but with a variation in the degree of protection. Further, the data on the effect of  $\text{FeSO}_4$  provided a clear picture about the role of  $\text{Fe}^{2+}$  in the enhancement of lipid peroxidation in presence and absence of MPG. It was also demonstrated that both the

spontaneous and radiation-induced lipid peroxidation of erythrocytes was increased by  $\text{FeSO}_4$ . The results also showed the enhancement of radiation-induced lipid peroxidation by MPG in the presence of  $\text{FeSO}_4$ . Such complex formation may be responsible for the lesser protective effect of MPG (DRF = 1.4) in mice. Further work needs to be done in *in vivo* systems.

Some discrepancies on WR-2721 also indicate that caution has to be adopted in the understanding of conditions in which thiol compounds might protect or fail to protect tumours (Lunec *et al.*, 1981). Reports on failure of WR-2721 to exhibit differential radioprotection of normal and cancerous tissues have also appeared (Rojas and Stewart, 1980; Rojas *et al.*, 1982; Clement and Johnson, 1982; Milas *et al.*, 1982). Philips *et al.*, (1973) had also failed to substantiate the very low protective effect on tumours by WR-2721 reported by Yuhas (1972, 1973) and had shown a wider variation in protection afforded to normal tissues than was previously reported. The importance of dephosphorylation of WR-2721 in the protective efficiency of the drug has been demonstrated by us in erythrocytes (Ayene and Srivastava, 1989; Srivastava, 1990). Protective and non-protective effect of WR-2721 in microsomes and erythrocytes have been indicated. The effect in erythrocytes was suggested because of the inadequate amount of dephosphorylating enzymes. This was confirmed by the addition of microsomes to erythrocytes that reduced the radiation damage of the erythrocytes. Such dephosphorylation has also been demonstrated in the presence of dephosphorylating enzyme, alkaline phosphatase (Misra and Srivastava, 1981; Collobro Jones *et al.*, 1985; Misra, Ayene and Srivastava, 1990a, 1990b). Research has also been carried out to study and regulate the metabolism, toxicity and radioprotection of WR-2721 by an inhibitor of alkaline phosphatase, levamisole (Brown *et al.*, 1986). Research in the direction of inhibiting the alkaline phosphatase activity in tumour cells may further enhance the chances of using WR-2721 in the differential protection of normal against the cancer cells.

### RADIOSENSITIZERS

Just like radioprotectors, sensitizers too are of two types, apparent and true. Radiosensitizers are drugs which have the capacity to increase the lethal effects of radiation. If they are non-toxic, then they will be ideal since they could kill the cancerous cells at lower doses of radiation without too much harmful effects on normal cells.

There are many apparent sensitizers such as Atinomycin D which depresses DNA dependent RNA synthesis; Puromycin which selectively depresses protein synthesis; Methotrexate which interferes with the synthesis of DNA by preventing the formation of coenzyme necessary for *in vivo* synthesis of thymidine; or 5-Fluorouracil which kills cells during the DNA synthetic or S phase of their cell cycle. They are not true sensitizers since they are not compounds which act only additively with radiation.

True radiosensitizers are drugs such as 5-Chlorodeoxyuridine, 5-Bromodeoxyuridine and 5-Iododeoxyuridine. The basis of their action depends on the fact that the combining size (van der Waals radius) of the atom of chlorine, bromine or iodine is very similar to the methyl group, CH<sub>3</sub>. The halogenated pyrimidines as they are called therefore replace the normal DNA precursor thymidine and hence weaken the backbone which breaks during division. Their use, however, been limited because of many limitations and also because they get metabolized very soon and hence become ineffective.

### SENSITIZERS OF HYPOXIC CELLS

Many compounds have been developed such as Metronidazole, Misonidazole, Nitroimidazole, etc., which are effective against specific hypoxic cells of tumours since they metabolise slowly and reach the hypoxic cells easily. These drugs increase the oxygen enhancement ratio of the cells upto 3.0. These compounds are being used clinically and have been successful to some extent.

### HYPERTHERMIA

The use of hyperthermia alone or in combination with radiation had been suggested over the years but it has found a permanent place in the management of cancer only for about ten years. In India, unfortunately, it has still to find a place. Miller *et al.* (1977) and Dewey *et al.* (1977) have reviewed the early clinical use of hyperthermia. More recently, Arcangeli *et al.* (1988) have published a review based on the works of seventeen groups of scientists and clinicians. They have described the effect of adjuvant hyperthermia on the radiation response. It has been shown that the frequency of response varies and the tumour regression can be enhanced from 25% to 65% depending upon the tumour and its site. The thermal enhancement ratio of radiation plus hyperthermia and radiation alone comes to 1.4 to 2.0. The degree of sensitization by hyperthermia and

radiation as compared to radiation alone varies from 1.5 to 4.3 depending upon the temperature used from 41°C to 43°C.

A few months back (about February-March 1990) Dr. Kenneth Alonso and his group of the University of Atlanta Medical Hospital have treated AIDS patients and the blood of patients have been shown to be HIV negative for more than four months. The blood of the patient was passed through a tube the temperature of which was raised to about 43°C. Further, terminal patients are being treated by this method. The doctors do not claim to have checked AIDS but are investigating the reasons for the blood remaining HIV negative for such a long time (Personal communication).

### IMMUNOTHERAPY

The growth of malignant tumour does not depend on cancer itself but it also depends on the body's response to it. If somehow or the other the body's immune response could be bolstered, then he or she may be able to fight off cancer better. Unfortunately, it has been easier said than done. The situation, hopefully, seems to be changing. In early seventies it was shown that BCG vaccine injections could induce guinea pigs' tumour to regress. Overzealous reports hailed BCG as a "Cancer Cure", which it was not. But the event did give fruitful leads.

Old therapies focussed on non-specific stimulators of immune system including BCG but more recently several compounds such as lymphokines, cytokines, interferons and inter-leukins have been developed which are anti-viral proteins- glyco-proteins- and are especially useful for cancers of viral origin. The new therapies use tumour cells, often prepared from patient's own concern, to elicit specific immune responses to the particular tumours from which the patients suffer.

Development of several vaccines are in the pipeline. Although vaccines do seem to prevent metastases, enough time has not elapsed since the study began to establish whether it improves 5-Year Survival ratio which are considered to be "gold standards" for successful cancer therapy. Although many are now optimistic about vaccine results, but people are cautious not to claim them as a cure for cancer at this stage. However, people had been trying to develop cancer immunotherapy for twenty to thirty years which has only recently started giving encouraging results.



## TUMOUR ANGIOGENESIS FACTOR

In early seventies, an observation was made that when tumour grows in size, blood capillaries start proliferating growing nearer and nearer to the tumour. This was a reaction to the growth of tumour since the blood capillaries can supply oxygen and nutrition upto only approximately 150  $\mu\text{m}$ . A pharmaceutical firm donated 23 million US Dollars to Harvard Medical School to carry on research to find out if any tumour angiogenesis factor was involved in this. After 12-13 years of work, when the scientists were on the verge of giving up the pursuit, they found some clue when they got a compound which they named as Angionin in 1985. They found 1 mg of the compound in 2000 litres of fluid in which cancer cells were grown. The molecular weight of angionin is 14,400 and has 123 amino acids and is 35% identical to ribonuclease. Biotechnology can now play a role to develop "anti-angionin factor".

## TUMOUR NECROTIC FACTOR

Lymphocytes and macrophages are known to produce a substance called Cachectin which is a tumour necrotic factor that helps in the process of necrosis of cancerous cells. It is a polypeptide and may be in dimeric, trimeric or pentameric forms. The molecular weight of each sub-unit is 17,000. It was sequenced in 1986 and has now been produced biotechnically. Numerous clinical tests are due in progress.

It had also been observed long time back that a cancer patient reacts to some drug and recovers but later on when a relapse takes place, the same patient does not respond to the same drug to which he/she had responded earlier. Scientists had been busy to find out the cause of such a phenomenon. It has again been shown very recently in 1986 that as a reaction to the drug, the cells produce certain glycoproteins whose molecular weight has been determined to be 170,000. This glycoprotein attaches itself with the membrane of cellular membrane. This changed cellular membrane now acts as a pumping mechanism and the next time the patient is given the same drug, it is pumped out of the cells before it could act on the cells.

## PHOTOCHEMOTHERAPY

Some pharmaceutical firms are now turning to photo-activated chemicals to formulate highly specific weapons against a variety of serious diseases including cancer. The first successes in this area were achieved by the

scientists of the American firm Johnson and Johnson, who developed 'Photofrin', a derivative of natural substances called porphyrins which are normally harmless components of haemoglobin in red blood cells. When exposed to sunlight, porphyrins emit an intense red fluorescence that produce ozone which destroys cells. Since 1985, another American pharmaceutical company, Cyanamid, began clinical testing of the drug on cancer patients at more than fifty centres.

What makes these drugs so effective are lasers and fibre-optic probes which deliver specific wavelengths of light needed to set off the drugs directly on cancerous tissue. For this, the drug is first injected into the patient where it collects in tumour cells, since they retain it more easily than healthy cells. After two days, the drugs are excreted out from the healthy cells and are retained only by the cancerous cells. A beam of low-powered laser light is channelled through a fibre-optic probe which activates the drug and selectively kills the cancerous cells, the normal cells remaining unaffected.

The origin of this development also lies in the past. About 100 years back, a Danish physician, Niels Fursen, used light from an arc lamp to treat a tubercular condition of the skin called lupus vulgaris. This disease was very common in northern latitudes, especially during winter months. Because of Finsen's discovery, lupus vulgaris practically disappeared from Scandinavian countries. Fursen's treatment was so original and successful that he was awarded one of the first Nobel prizes in 1903.

### CONCLUSION

In conclusion this can only be said that the fight for the cure of cancer is still on and will perhaps continue for some time to come. In general, one third of cancers are preventable, one third of cancers are curable and the remaining one third of cancers at present can only be treated with pain relief measures.

In India, approximately 50% cancers belong to oral cancer, whose detection in early stages is very much easier than many other cancers and hence are totally or nearly curable and cervical cancer which takes five years to develop localised cancer and another ten years to reach the advanced stage. The latter can again be detected early and if proper care is taken it can virtually be cured 100%. Thus, in India, preventive oncology should play a big role.

In the end, I will like to thank the Indian National Science Academy for awarding me this prestigious Aryabhata Memorial Medal.

## References

1. Archangeli G, Overgaard J, Gonzalez DG, and Srivastava PN 1988 Hyperthermia trials; *Int J Radiation Oncology Biol Phys* 14 S93-S109.
2. Ayene SI, Kale RK and Srivastava PN 1988 Radioprotective effect of 2-mercaptopropionylglycine on radiation-induced peroxidation and enzyme release in erythrocytes; *Int J Radiat Biol* 53 629-639.
3. Ayene SI and Srivastava PN 1985 Radioprotective effect of 2-mercaptopropionylglycine on radiation-induced microsomal lipid peroxidation; *Int J Radiat Biol* 48 197-205.
4. Ayene SI and Srivastava PN 1989 Effect of WR-2721 on radiation-induced lipid peroxidation and enzyme release in erythrocytes and microsomes; *Int J Radiat Biol* 56 265-275
5. Brown DQ, Shaw LM, Pittock JW, Mann DJ, Hardiman J, Pogach, R and Gold J 1986 Modification of WR-2721 toxicity and radioprotection by an inhibitor of alkaline phosphatase; *Int J Radiation Oncology Biol Phys* 12 1491-1493
6. Clement JJ and Johnson RK 1982 Influence of WR-2721 on the efficacy of radiotherapy and chemotherapy of murine tumours, *Int J Radiation Oncology Biol Phys* 8 539-542.
7. Collobro Jones PM, Fahey RC, Smoluk GD and Ward JF 1985 Alkaline phosphate promotes radioprotection and accumulation of WR-1065 in V79-171 cells incubated in medium containing WR-2721; *Int J Radiat Biol* 47 23-27.
8. Dev PK, Pareek BP, Gupta SM, Goyal PK and Mehta G 1982 MPG (2-mercaptopropionylglycine) protection against growth inhibiting effects *in utero* irradiated mice; XIII Annual Meeting of the Radiation Research Society, USA. Abstract.
9. Dewey WC, Hopwood LE, Sapareto SA and Gerweck LE 1977 Cellular responses to combinations of hyperthermia and radiation; *Radiology* 123 463-474.
10. Lunec J Cullen B Walker H and Hornsey SA 1981 Cautionary note on the use of thiol compounds to protect normal tissues in radiotherapy; *Brit J Radiol* 54 428-429.
11. Milas L, Himder N and Reid BO 1982 Protective effects of WR-2721 against radiation-induced injury on murine gut, testis, lung and lung tumour modules; *Int J Radiation Oncology Biol Phys* 8 535-538.

12. Misra VS, Ayene SI and Srivastava PN 1990a Radiolysis of alkaline and acid phosphatase in presence and absence of MPG and WR-2721: in aqueous solution; Communicated.
13. Misra VS, Ayene SI and Srivastava PN 1990b Radiolysis of alkaline and acid phosphatase in presence and absence of MPG and WR-2721: in whole homogenates; Communicated.
14. Misra VS, and Srivastava PN 1981 Radioprotective effect of MPG and WR-2721 against gamma-radiolysis of human placental alkaline phosphatase; *Proc natn Acad Sci (India)* **51B** 318-324.
15. Phillips TL, Kane L and Utley JF 1973 Radioprotection of tumour and normal tissues by thiophosphate compounds; *Cancer* **32** 523-535.
16. Rojas AM and Stewart FA 1980 Radioprotection by WR-2721 in normal mouse skin and tumour; *British Institute of Radiology: Radiology Work-in-Progress Meeting* Radiosensitisers and Radioprotectors. November 21.
17. Rojas AM, Stewart FA and Denekamp J 1982 Experimental radiotherapy with WR-2721 and misonidazole; *Int J Radiation Oncology Biol Phys* **8** 527-530.
18. Saharan BR, Saini MR and Uma Devi P 1978 MPG protection and goblet cell kinetics in mouse jejunum, *Strahlentherapie* **154** 60-62.
19. Saharn BR and Uma Devi P 1977 Radiation protection of mouse testes with 2-mercaptopropionylglycine; *J Radiat Res* **18** 308-16.
20. Saini MR and Uma Devi P 1979a MPG (2-mercaptopropionylglycine) protection in thymus and spleen against 1500 R gamma radiation; *Ind J. Expt Biol* **17** 942-952.
21. Saini MR and Uma Devi P 1979b Modification of radiation injury in thymus of Swiss albino mice; *Ind J Expt Biol* **17** 782-783.
22. Saini MR, Uma Devi P and Yadav SS 1978 Radiation protection of bone marrow lymphocytes by 2-mercaptopropionylglycine (MPG); *Experientia* **34** 1627-1628.
23. Srivastava PN 1990 Sulphur compounds and chemical radioprotection; in *Recent Trends in Radiobiological Research* ed P Uma Devi (Jodhpur, India, Scientific Publishers, 5-11).
24. Sugahara T and Srivastava PN 1976 2-Mercaptopropionylglycine (MPG): A review on its protective action against ionizing radiations. In *Modifications of Radiosensitivity of Biological Systems* (Vienna: International Atomic Energy Agency), 77-87
25. Uma Devi P 1977 Protection of mouse intestine against gamma-irradiation by 2-mercaptopropionylglycine *J Radiat Res* **18** 160-163.
26. Uma Devi P and Saini MR 1977 Protection of mouse thymus against cobalt 60 radiation by 2-mercaptopropionylglycine (MPG); *J Radiat Res* **18** 211-224.

27. Uma Devi P, Saini MR, Verma A and Saharan BR 1978 Radioprotective effect of 2-mercaptopropionylglycine on the small intestine of Swiss albino mice; *Int J Expt. Biol.* 16 86-88.
28. Uma Devi P and Saharan BR 1978 Chemical protection of mouse spermatocytes against gamma rays with 2-Mercaptopropionylglycine; *Experimentia* 34 91-92.
29. Yuhas JM 1972 Improvement of lung-tumour radiotherapy through differential chemo-protection of normal tumour tissue; *J Nat Cancer Inst* 48 125-1257.
30. Yuhas JM 1973 Radiotherapy of experimental lung tumours in the presence of a radioprotective drug, S-2-(3-aminopropylamino) ethylphosphorothioic acid (WR-2721); *J Nat Cancer Inst* 50 69-78.



**Obaid Siddiqi** (b. 7 January 1932) did Ph.D. (1961) from University of Glasgow, UK and was awarded D.Sc. (h.c.) from Aligarh Muslim University (1984) and Banaras Hindu University, Varanasi (1991). He is Eminent Professor, Tata Institute of Fundamental Research, Bombay.

After some early research in plant embryology and wheat genetics, Siddiqi mapped the fine-structure of the Paba gene of *Aspergillus*, while at Glasgow. His work provided a strong evidence for polarized intragenic recombination. At the University of Pennsylvania, Siddiqi and Garen discovered the

suppressors of 'nonsense' mutations and this work led to the discovery, at TIFR, of 'nonsense' codons, the stop signals in the genetic code. He and his associates were able to show that DNA transfer can be dissociated from replication, and that recombinant DNA molecules can arise from conserved unreplicated DNA. He has also done much work on behavioural genetics and neurobiology of *Drosophila*. He and his associates have done pioneering work on neurogenetics of the chemical sense of *Drosophila*, identifying a series of genes whose mutations block olfactory or gustatory responses. Their work has opened up the prospect of an integrated genetic and neurobiological investigation of chemosensory perception. He has contributed a chapter on "Alleles" to 'McGraw-Hill Encyclopedia on Science'; authored 'Mechanisms of Gene Action and Regulation - The Biological Basis of Medicine' Academic Press.

Siddiqi is Fellow of The Royal Society, London, Indian Academy of Sciences (Vice-President 1983-86, President 1986-89), Maharashtra Academy of Sciences, Third World Academy of Sciences, and National Academy of Sciences (India) has been Member of INSA Council (1981-83). He is the recipient of UGC National Lecturer (1974); Shanti Swarup Bhatnagar Prize (1976), Golden Jubilee Commemoration Medal (INSA) (1986); Birla Smarak Kosh National Award (1989); Goyal Prize (1991); Aryabhata Medal (INSA) (1992), Padma Bhushan (1984).

*Obaid Siddiqi was elected to the fellowship of the Academy in 1977.*

**Address:** Professor, Molecular Biology Unit, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005; Tel.:215 2971/79; Telex: 011-83009 Code: TIFR IN; Fax. 091-22-2152110.

# **MOLECULAR BIOLOGY OF LEARNING AND MEMORY**

O SIDDIQI

I wish to express my gratitude to the Indian National Science Academy for giving me the Aryabhata award. Aryabhata the famous astronomer and mathematician (476 AD) has been described as the creator of Indian Astronomy. To be connected with his name, howsoever indirectly (and undeservingly) at this great distance in time, is indeed a great honour. I take this opportunity to thank the council of INSA and its President for their kindness.

The topic of my lecture, learning and memory is a subject which has interested speculative minds in all ages. Aryabhata too may have thought about it. Aristotle and St. Thomas Aquinas certainly did. It still remains a largely unresolved subject. Some recent developments in molecular neurobiology promise to provide us with insights into biochemical mechanisms of learning. It is to these that I am going to draw your attention.

Learning involves acquisition of new knowledge. Memory has to do with storage of acquired knowledge. Both imply a change in behaviour with experience. This definition is very broad and includes a hierarchy of phenomenon.

Sensitization

Habituation

Associative conditioning (Pavlovian learning)

Higher learning involving cognition and awareness

Is there a common elementary mechanism underlying these very diverse processes? Simplest organisms are capable of elementary forms of learning but, properly speaking, learning requires a certain level of complexity in the system, in particular a set of interconnected neurons. D.O. Hebb suggested some fifty years ago that in the course of learning, connections between neurons undergo changes.

This change in synaptic connections or "synaptic efficacy" is the key to the problem of the memory trace.

## MOLECULAR MECHANISMS

The first attempts to formulate an explicit molecular theory of learning date to early days of molecular biology. It was suggested that learnt information could be stored in RNA. The claim was that, when trained *Planaria* (flat worms) are mashed and fed to their naive brothers, acquired behaviour is transmitted to cannibals. The vehicle of memory transfer, it was claimed, were the RNA molecules. These irreproducible experiments based on ill-conceived theories soon fell by the wayside.

Real insights into molecular mechanisms of synaptic changes came with the discovery that proteins undergo chemical modifications such as phosphorylation of certain amino acids. These modifications are rapid and do not require *de novo* protein synthesis. Activities of protein molecules can thus be changed without recourse to genetic mechanisms. Protein modifications such as phosphorylation or methylation are controlled by small molecules like cyclic AMP, the so called 'second messengers'. Second messengers in cellular signalling provide a plausible mechanism for triggering short term memory in nerve cells.

Certain hormones and neurotransmitters bind to receptors on the cell membrane and activate an enzyme adenylate cyclase which converts ATP to cyclic AMP. Cyclic AMP activates a second set of enzymes, the *kinases* which can phosphorylate proteins. Phosphorylation either increases or decreases the biological activity of target proteins. Cyclic AMP is one of the second messengers. There are other such as IP<sub>3</sub> (inositol 1,4,5, - triphosphate), diacylglycerol or Ca<sup>2+</sup> which act through other biochemical pathways to connect external signals to persistent internal changes in cellular biochemistry. (Fig. 1)

Phosphorylation of membrane proteins can change the signalling or computational properties of neurons and alter presynaptic and post-synaptic efficacy. There are two major classes of target proteins whose modifications affect synapses; channels which allow the inflow and outflow of ions such as Na<sup>+</sup>, K<sup>+</sup>, Ca<sup>2+</sup> and Cl<sup>-</sup> and receptor-coupled enzymes which are activated by second messengers.

## ASSOCIATIVE CONDITIONING

Ivan Pavlov, the Russian physiologist, found that when a conditioned stimulus (CS) is paired with an unconditioned stimulus (UCS) repeatedly, the response initially evoked by UCS, can be evoked by CS. For conditioning to occur, CS and UCS must be paired closely in time.



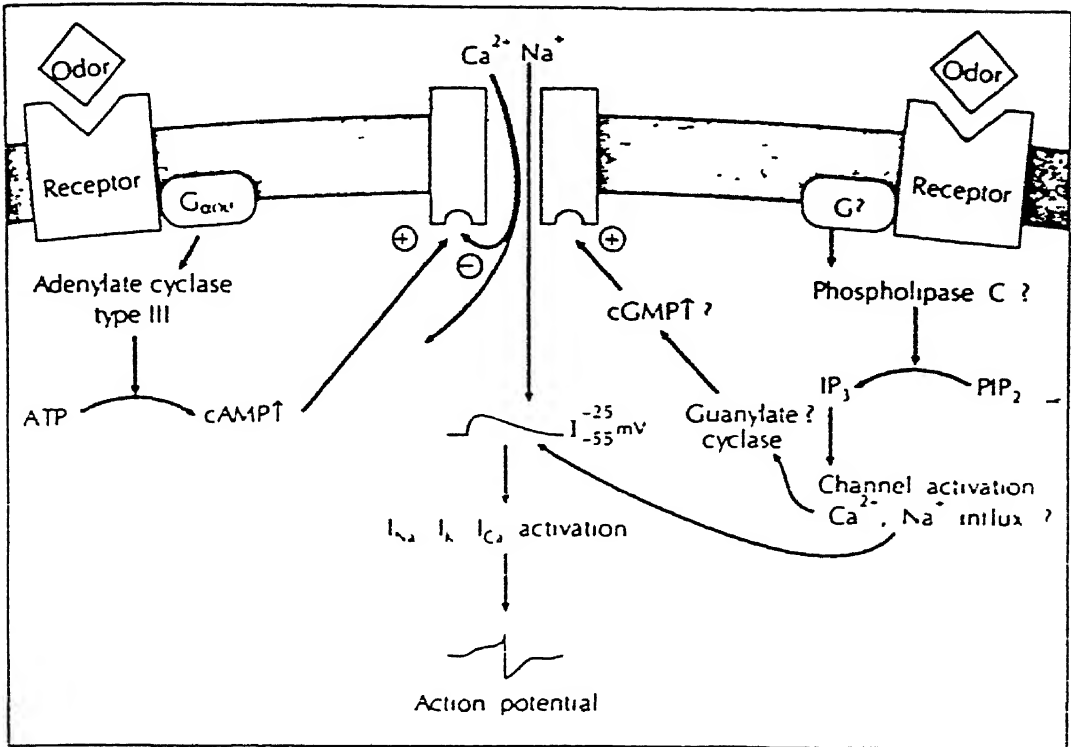


FIG 1 Excitation of nerve cells by external stimuli involves a variety of signalling systems. The diagram summarises recent work on olfactory sensory neurons. A number of second messengers, cAMP, IP<sub>3</sub> and Ca<sup>2+</sup> can activate kinases which phosphorylate synaptic proteins. Protein modifications accompanying electrical activity of neurons are the basis of short term memory.

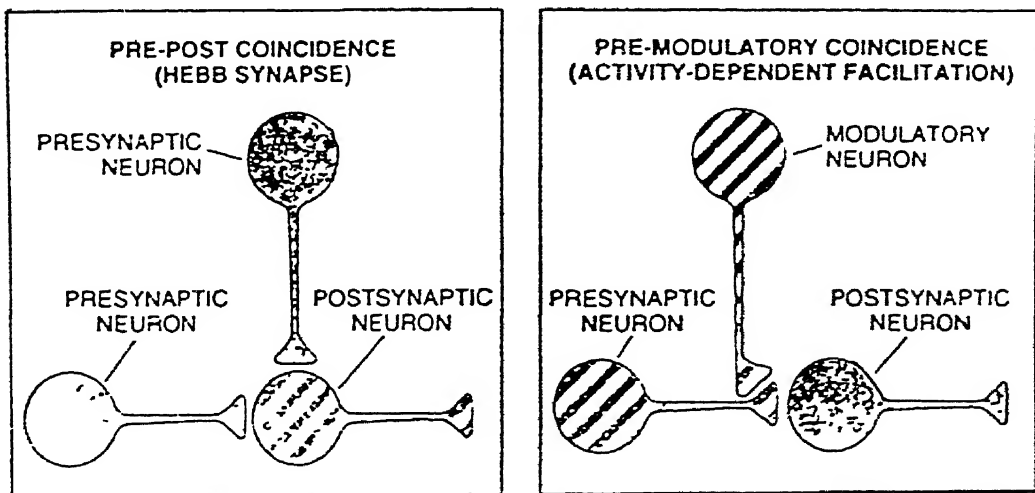


FIG 2 Associative memory in neurons arises by strengthening of synaptic connection by A. Coincident activity in the presynaptic and post-synaptic cell (Hebb's hypothesis) B. Coincident activity of a modulatory neuron synapsing with a presynaptic terminal (Tauc-Kandel model).

Classical Pavlovian conditioning according to Hebb's interpretation requires coincident activity in the presynaptic and the post synaptic neuron. When this coincidence is repeated the synapse is strengthened. In 1963, Kandel and Tauc discovered a new associative rule. They found that connections between two neurons could be strengthened with the help of a third modulatory neuron. The modulatory neuron in the Kandel-Tauc model acts on the presynaptic terminal. Coincident activity in the presynaptic neuron and the modulatory neuron changes the strength of the synapse. The two situations are diagrammed in Fig 2.

### GILL WITHDRAWAL REFLEX IN *APLYSIA*

The molecular mechanism of synaptic strengthening was first worked out by Kandel and his associates, working with the sea hare *Aplysia californica*. *Aplysia* has about 20,000 large neurons which can be recognized individually and impaled with electrodes. This permits a detailed analysis of the neural circuits which underlie various aspects of *Aplysia's* behaviour. One of these behaviours, much studied by physiologists, is the gill-withdrawal reflex stimulated by touching the siphon or the mantle shelf. The response (CS) can be strongly conditioned by pairing it with an electric shock to the tail (UCS). About five trials are enough to induce learning. The behavioural circuit is outlined in Fig. 3. A modulatory neuron releases serotonin which acts on the presynaptic terminal of the sensory neuron. Serotonin released by the UCS activates adenylate cyclase in the presynaptic terminal of the sensory neuron, cyclic AMP in turn activates a protein kinase which phosphorylates  $K^+$  channels to cause prolonged depolarization and calcium influx. Incoming calcium binds to calcium calmodulin and further activates adenylate cyclase - a kind of positive feed back. Through a second pathway involving phospholipase and protein kinase C, synaptic vesicles are mobilized to release neurotransmitter at the motor neuron. This, somewhat complicated biochemistry leads to increased synaptic efficacy. A second stimulus to the sensory neurons of the siphon elicits a stronger motor response thanks to protein phosphorylations caused by previous experience.

The fruitfly, *Drosophila melanogaster* can be trained to avoid specific odours by punishment with electric shock. W.G. Quinn isolated a number of mutants with impaired learning and memory. These mutations have been shown to affect various steps in cAMP-dependent phosphorylation. In the mutant called *dunce* the enzyme

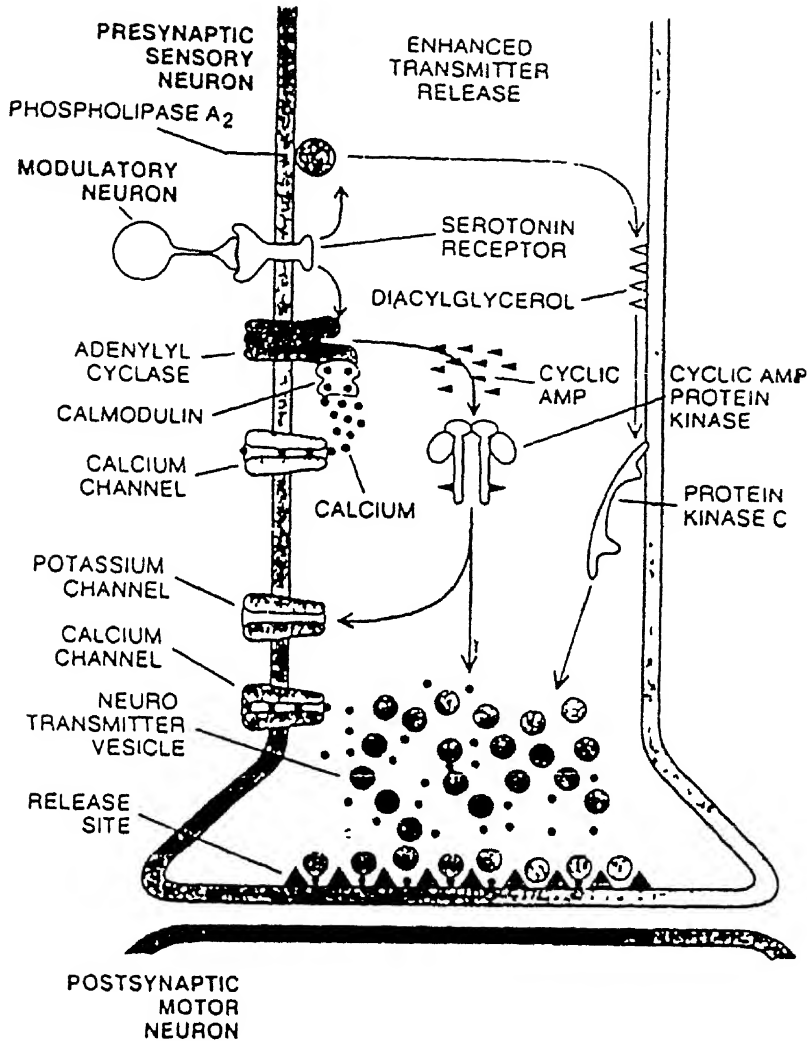


FIG 3 Biochemistry of associative conditioning in the gill withdrawal reflex of *Aplysia*. The release of serotonin by a presynaptic modulatory neuron starts a chain of signals leading to protein modifications which increases the efficacy of the synapse between the sensory neuron and the motor neuron.

phosphodiesterase is defective. The dunce flies are unable to consolidate long term memory. In the mutant *rutabaga* a calcium-calmodulin dependent adenylyl cyclase is affected. It has been shown recently that the enzymes of phosphorylation metabolism involved in the learning pathway are concentrated in the mushroom body, a part of the fly's brain which is believed to be the association centre. Analysis of learning mutants of *Drosophila* thus leads us to the same biochemical mechanisms that are revealed by neurophysiological experiments on the gill withdrawal reflex of *Aplysia*. (Fig. 4)

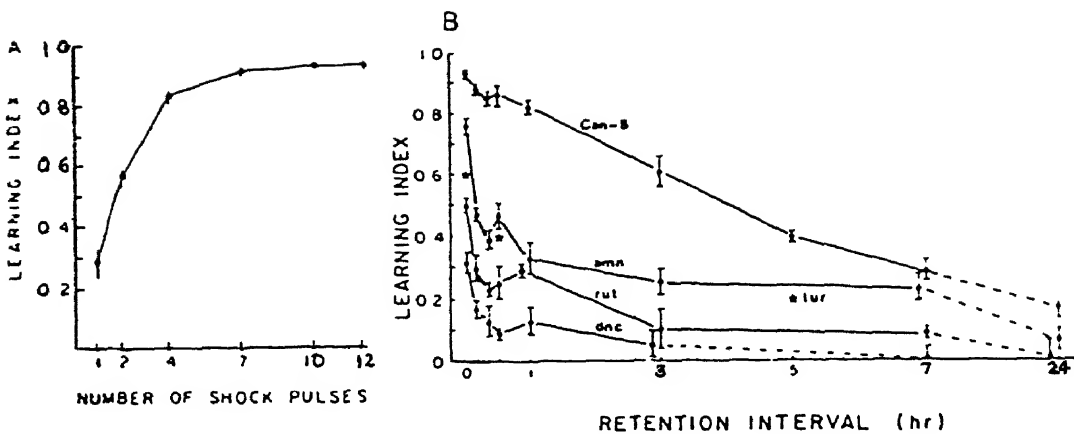


FIG 4 Learning mutants of *Drosophila*. Left. The flies can be conditioned to avoid particular odours by means of electric shock. Right. Genetic mutants, *Dunce*, *rutabaga*, *cabbage* and *amnesiac* are unable to convert short term memory into long term memory. Some of these genes control the pathways of phosphorylation cascade.

### HEBBIAN SYNAPSE AND NMDA-RECEPTORS

Synaptic strengthening of the type postulated by D.O. Hebb has been demonstrated in the hippocampus of vertebrates. Hippocampus is a temporary repository of visual memory in its passage from sensory to pre-frontal cortex. Neurons of hippocampus display plastic behaviour characteristic of learning. Stimulated electrical activity in the hippocampal pathways leads to increased synaptic efficacy. The phenomenon is called long term potentiation (LTP), it can be studied *in vitro* in slices of hippocampus and is currently the object of intensive world wide research. Associativity in the hippocampal pathways is of the classical Hebbian type, a presynaptic to post-synaptic strengthening. Its molecular basis is somewhat different from that of Pavlovian conditioning. LTP is mediated by glutamate and N-methyl-D-aspartate (NMDA) receptors. The NMDA channel is blocked in the resting state by  $Mg^{++}$ . If the neuron is depolarised, let us say, by the activation of some non-NMDA channel, the NMDA channel is unblocked leading to an influx of calcium and the activation of the second messenger system. NMDA channels thus act as coincidence detectors with properties analogous to adenylate cyclase. There is an important difference. Induction of LTP depends upon post-synaptic calcium influx but the maintenance of LTP requires an enhanced presynaptic release of transmitter. A signal must, therefore, go from the

post synaptic to the presynaptic terminal. The messenger for this retrograde signal is believed to be nitric oxide.

### HIGHER LEARNING

In 1950, Brenda Milner described a remarkable case of amnesia in a patient after bilateral removal of temporal lobes. The patients lost the ability to form new long term memories although he retained previously acquired memories. This kind of memory loss has been extensively investigated in men and monkeys by Mishkin and others. These studies throw light on the role of temporal lobe in memory formation.

Visual memory formation in primates starts from the striate cortex, where a topographic map of the visual field is laid down. This is the so called 'primal sketch' of David Marr. A more global representation of the visual scene is then transmitted to the inferior temporal cortex whose role has been greatly clarified by experiments involving temporal lesions. There seem to be two parallel pathways running through hippocampus, amygdala and diencephalon. Bilateral lesions of hippocampus and amygdala cause global amnesia. If hippocampus alone is damaged 'spatial memory' is affected but 'object memory' remains intact. Amygdala on the other hand is critically necessary for associating visual memory with memories in other modalities. We have not even begun to understand the detailed mechanisms of global memory but proteins whose synthesis and modifications are correlated with memory formation in the brain of higher animals are at present, the object of active research. One such protein is called F1. It has been shown that phosphorylation of F1 by protein kinase C follows the pathway of visual memory in the temporal lobes.

### CONCLUSION

Short term changes in synaptic efficacy are accompanied with molecular modifications in proteins. Phosphorylation has a pronounced effect on the activity of synaptic proteins. Different form of learning and memory, associative learning, habit formation or cognitive learning may employ somewhat different pathways but the elementary biochemical mechanisms are probably universal.

Maintenance of memories over long durations requires other, stable alterations in the wiring diagram of the brain, number of functional receptors, number and location of synapses or size and extent of neuronal arborizations. The molecular mechanisms that I have discussed have the

attractive feature that they allow us to connect short term changes with long term changes in a natural way through gene activation. At any rate we seem to have come a long way from the neuropsychology of yesteryears which looked upon cortex as an ill-defined tissue and the memory trace as an enigma.



**Sudhanshu Shekhar Jha** (b. 25 December 1940) Ph.D.(1965) from Stanford University, USA. He is Distinguished Professor (1990- ) and was formerly Reader, Associate Professor, Professor, Senior Professor (1958-90) at Tata Institute of Fundamental Research, Bombay

Jha's major research achievements relate to (i) pioneering theoretical predictions of optical second-order nonlinearity at metal surfaces and in centrosymmetric plasmas, (ii) calculation of optical nonlinear susceptibilities of important doped semiconductors, (iii) a technique for calculating energy

distribution of thermal neutrons in a finite moderator assembly, (iv) a theory for the surface-enhanced Raman scattering from adsorbed molecules at metal surfaces, (v) superconductivity due to exciton-exchange and electronic mechanisms in metal-semiconductor sandwich structures and in highly correlated high- $T_c$  layered metals, and (vi) ultra-fast nonequilibrium optoelectronic devices and transient electrical response of photo-excited carriers in semiconductors. He is the Guest editor/joint editor of. 'Recent Advances in Raman Spectroscopy' (World Scientific, 1989).

Jha is Fellow of Indian Academy of Sciences, National Academy of Sciences (India), and American Physical Society. He is the recipient of Shanti Swarup Bhatnagar Prize (1979); C.V. Raman Centenary Medal (Indian Association for the Cultivation of Science) (1988); Jawaharlal Nehru Birth Centenary Lecture Award (INSA) (1991), S.N. Bose Medal (INSA) (1992)

*Sudhanshu Shekhar Jha was elected to the fellowship of the Academy in 1979*

**Address:** Distinguished Professor, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay 400 005, Tel. (022) 2152971, Telex 011 83009 TIFR IN; Fax: 91-22-2152110, E-mail: ssjha@tifrvax.bitnet; Res 602, Meghnad, TIFR Housing Complex, Homi Bhabha Road, Colaba, Bombay 400 005; Tel: (022) 2152184.

# FROM OPTICS TO OPTOELECTRONICS AND PHOTONICS

SUDHANSHU S JHA FNA

*A short review of the progress made in optics towards the development of the emerging field of optoelectronics is presented. It is argued that the new field is not only changing the face of information technology, at present based mostly on semiconductor microelectronics, but it is also making it possible to examine experimentally many fundamental questions in basic physics, at the frontiers of knowledge. The march of optics towards the technology of pure "photonics" is also discussed briefly*

## INTRODUCTION

Optics has been an important field of science for a very long time<sup>1</sup>. From the very beginning, great philosophers and thinkers have tried to explain the nature of light, its emission and transmission through different kinds of matter, formation of images, magnification by lenses, colour dispersion, and the process of vision. Starting from the great work of Hero during the first century AD on the straight transmission of light rays, we must mention the names of Ptolemy, Alhazen, Galileo, Kepler, Descartes, Snell, Fermat, Grimaldi, Huygens, Newton, Young, Fresnel, Arago, Fizeau and Foucault who made important contributions to optics till the late nineteenth century. Eventually, these developments culminated in the discovery of the extremely elegant electromagnetic wave theory of light by Maxwell, which was verified experimentally, soon after his death, by Hertz in 1887. Maxwell's electromagnetic theory remains a strong pillar for the present day physics.

It is well known that the fabrication of optical telescope by Galileo led to the rapid development of modern astronomy and astrophysics, whereas the discovery of optical microscope revolutionized the study of medicine and biology. But much more than that, it can be argued that all major revolutionary concepts introduced in modern physics of to-day owe their origin to optics. The Michelson-Morely experiment to determine the constancy of the speed of light, the determination of the detailed characteristics of the photoelectric effect, the analysis of line spectra from atoms and molecules, the observation of the bending of light in strong gravitational field, and accurate measurements in optical spectroscopy



during the early twentieth century were crucial in establishing the basic foundations of new physics in the form of the special theory of relativity, the concept of light quanta (photons), the old quantum theory, the general theory of relativity and the modern quantum mechanics and electrodynamics.

Since human eyes are very good optical detectors, although quite slow in resolving images (a few per second), and since our sun is a very powerful source of light, optics has always been very useful in everyday life. Storage of information and its propagation *via* handwritten manuscripts and printed books or through photographs taken by a camera, reading and normal vision through optical lenses, and navigation and signalling through light houses, etc., are some of the familiar examples. The utility of telescopes, microscopes, cameras and many other optical instruments need not be overemphasized here. But, as far as the recent progress in the field during the last few decades is concerned, the 1960-discovery of laser as an intense coherent source of light has turned out to be a real triumph of optics and optical spectroscopy. By now, tunable and powerful oscillators and amplifiers are available in almost the entire range of infrared and visible frequencies, which can be pulsed up to the temporal width of a few femtoseconds ( $10^{-15}$ s), modulated at the rate of  $10^{11}$  Hz or more, and transmitted through large distances *via* extremely low-loss optical glass-fibres or through free space. Because of very large bandwidth for the optical carrier wave, it has obvious advantages in the field of communications and informatics, provided that it can be modulated at extremely high rates. With the parallel development of extremely efficient, compact, and fast semiconductor detectors for optical photons, having response time as small as a few picoseconds ( $10^{-12}$ s), and the fact that the optical image processing is inherently two-dimensional, once again optics is helping us to solve existing problems in basic sciences and technology. It is not only being used to investigate important problems in chemistry, biology and other sciences, including extremely fast dynamical processes in matter<sup>2</sup>, but many crucial optical photon experiments are being done to detect gravitational waves<sup>3</sup> and to test the breakdown, if any, of the most fundamental theoretical concept of modern physics—quantum mechanics<sup>4</sup>! It has also given rise to the new hybrid technology of “Optoelectronics”, with the future aim of developing pure “Photonics”. These activities, both in pure science as well as in technology, are indeed at the frontiers of knowledge. Instead of presenting even a glimpse of all these developments, for obvious reasons, our limited aim here will be to discuss how modern optics is

leading us to the new technology of optoelectronics and photonics which will eventually replace the conventional microelectronics technology in many important applications.

In what follows, we will first introduce the basic principles of optics which are key to most of the technological developments. We will then discuss the present scope of optoelectronics. We will conclude this presentation with some remarks regarding the march towards photonics.

### OPTICS: BASIC PRINCIPLES

Three most basic physical processes involving light are its emission, absorption and scattering. The process of "vision" or optical detection of an object involves scattering of light emitted from a source by the object; the scattered wave being detected by a suitable optical detector, eg., eyes. The amplitude and phase of the scattered light contain enough information necessary to reconstruct the object, but the actual nature and quality of the image depends on the method and the type of detector being used for the purpose. For example, ordinary recording of images on photographic films, where only intensities of the scattered wave matter, differs considerably from holographic recording in which the complete phase information of the scattered wave is also kept as an interference pattern between the scattered wave and a reference beam. Light, of course, propagates in free space as a transverse wave, and two spatially coherent beams show the familiar interference phenomena. Interference of light arising from different points in the same source leads to the diffraction phenomena. When a light wave of wavelength  $\lambda$  is partially blocked by an opaque object of size  $D$ , or when it passes through a slit (an aperture of size  $D$  with spatially varying transmission properties), one gets the familiar far-field Fraunhofer diffraction pattern at distances  $d \gg D^2/\lambda$ , from the object plane. For  $d < D^2/\lambda$ , one gets the Fresnel diffraction pattern in the near-field region. Although, both patterns contain all the necessary information about the object, it is normal to use the more simple far-field pattern in recording images in holography and Fourier-transform (F.T.) optics<sup>1,5</sup>. In fact, the amplitude of the far-field pattern at the image plane is directly proportional to the two dimensional spatial Fourier transform  $G(f_x, f_y)$  of the optical signal  $g(X_o, Y_o)$  in the object plane. If  $d$  is the perpendicular distance of the image plane from the object plane, (i.e., difference between the Z-coordinates of the planes) and  $x_i, y_i$  are suitable cartesian coordinates in the image plane with the line joining the origins of the two planes being along the Z-axis, the amplitude at the point  $(x_i, y_i)$  gives F.T. Function

$G(f_x, f_y)$  with  $f_x = x_i/\lambda d$  and  $f_y = y_i/\lambda d$ . Thus, the ratio of the size of the recording image plane and the distance  $d$  determines the maximum of the spatial frequencies recorded, the reciprocal of which gives the minimum distance between two points in the object plane that can be resolved.

The propagation of light as well as all other classical characteristics of optical electric and magnetic fields  $\mathbf{E}(\mathbf{r}, t)$  and  $\mathbf{B}(\mathbf{r}, t)$  respectively, are determined by the solutions of Maxwell's equations

$$\nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0, \quad \nabla \cdot \mathbf{B} = 0 \quad \dots 1$$

$$\nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J}, \quad \nabla \cdot \mathbf{E} = 4\pi\rho \quad \dots 2$$

with given initial conditions. Note that the total current density,  $\mathbf{J}(\mathbf{r}, t)$  and the charge-density  $\rho(\mathbf{r}, t)$  satisfy the continuity equation

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0. \quad \dots 3$$

In the presence of a material medium, these equations are to be solved with the additional knowledge about the constitutive relations for  $\mathbf{J}$  and  $\rho$  as functions of  $\mathbf{E}$  and  $\mathbf{B}$  at each space-time point. If one introduces the generalized polarization field  $\mathbf{P}(\mathbf{r}, t)$  as

$$\mathbf{J} = \frac{\partial \mathbf{P}}{\partial t}, \quad \rho = -\nabla \cdot \mathbf{P}, \quad \mathbf{D}(\mathbf{r}, t) \equiv \mathbf{E}(\mathbf{r}, t) + 4\pi\mathbf{P}(\mathbf{r}, t) \quad \dots 4$$

it implies that we must equivalently know the constitutive relation for  $\mathbf{P}$  as a function of  $\mathbf{E}$  and  $\mathbf{B}$ .

At optical frequencies, one can very often use the long wavelength ( $\lambda \gg$  atomic size  $a$ ) electric-dipole approximation for finding the functional form of  $\mathbf{P}$ , which, in general, will be a nonlinear function of  $\mathbf{E}$ . In linear optics, the Cartesian components of  $\mathbf{P}$  at frequency  $\omega$  (i.e. temporal F.T. of  $\mathbf{P}_1^L$ ) are given by

$$P_i^L(\mathbf{r}, \omega) = \sum_{j=1}^3 \chi_{ij}^{(1)}(\omega) E_j(\mathbf{r}, \omega); \quad D_i^L = \sum_j \epsilon_{ij}^L E_j; \quad \epsilon_{ij}^L = 1 + 4\pi\chi_{ij}^{(1)} \quad \dots 5$$

For an isotropic or cubic material, the linear dielectric tensor  $\epsilon_{ij}^L$  is scalar, with  $\epsilon_{ij}^L(\omega) = \epsilon(\omega) \delta_{ij}$ . If one writes  $\epsilon(\omega) = (n + i\kappa)^2$ ,  $n(\omega)$  corresponds to the refractive index of the medium whereas  $\kappa(\omega) = c\alpha(\omega)/2\omega$  determines its absorption coefficient  $\alpha(\omega)$ . In a weakly absorbing medium, the phase velocity  $V_\phi = \omega/q \simeq c/n(\omega)$ , and the group velocity  $V_g = \partial \omega / \partial q$ , where the propagation vector  $q = [\omega n(\omega/c)] \hat{q}$ . In an anisotropic or non-cubic material,  $\epsilon_{ij}^L(\omega)$  is no longer a scalar, and waves in different directions can have different speeds, as studied in crystal optics. In an inhomogeneous medium, with characteristic length scale  $l$  for the spatial variation of  $\epsilon(\omega, \mathbf{r})$  large compared to the wave length of light  $\lambda$ , one uses the eikonal (ray) equations of geometric optics to determine optical propagation:

$$E(\mathbf{r}, t) = E_0 e^{i\Psi(\mathbf{r}, t)}, \quad \Psi(\mathbf{r}, t) \simeq -\omega t + \frac{\omega}{c} S(\omega, \mathbf{r}), \quad \dots (6)$$

$$\omega = -\partial \Psi / \partial t, \quad q(\mathbf{r}) = \nabla \Psi = \frac{\omega}{c} \nabla S(\omega, \mathbf{r}) \quad \dots (7)$$

and

$$|\nabla S|^2 = \epsilon(\omega, \mathbf{r}). \quad \dots (8)$$

The solution for the ray vector  $\nabla S(\omega, \mathbf{r})$ , in fact, traces out rays of light of frequency  $\omega$  to describe the laws of reflection, refraction and image formation, in the presence of optical elements like lenses, mirrors, prisms, etc., distributed in space. It can also describe the characteristics and resolving power for many usual optical elements like diffraction gratings and Fabry-Perot interferometers and filters.

Till now, we have been describing optical fields as classical  $c$ -number fields. In a homogeneous medium, with no boundaries, the electric field, of a light wave at frequency  $\omega$  can be expanded in terms of plane waves  $E_{q\mu} \hat{e}_{q\mu} \exp(i\mathbf{q} \cdot \mathbf{r} - i\omega t)$ , where  $q^2 = (\omega^2/c^2) \epsilon(\omega)$ , and where  $\hat{e}_{q\mu}$ ,  $\mu = 1, 2$ , are unit vectors for two transverse polarizations. In a bounded space with volume  $V$ , e.g., in dielectric waveguides, the ortho-normalized eigen-functions  $U_{m\mu}(\mathbf{r}) / \sqrt{V}$  satisfying the relations

$$\frac{1}{V} \int U_{m\mu}^*(\mathbf{r}) U_{m'\mu'}(\mathbf{r}) d^3r = \delta_{mm'} \delta_{\mu\mu'} \quad \dots (9)$$

are no longer plane waves. Thus, in general, one has the expansion

$$E = \sum_m \sum_{\mu} \left[ E_{m\mu}^{(\omega)} e_{m\mu} \frac{1}{\sqrt{V}} U_{m\mu}(r) e^{-i\omega t} + \text{C.C.} \right] \quad \dots (10)$$

The electromagnetic field energy for a given single-mode is then given by  $E(\omega)^2/2\pi$ . For a given single -mode, we may formally write

$$E^{(\omega)} = E_R^{(\omega)} + iE_i^{(\omega)}, \quad |E^{(\omega)}|^2 = E^{(\omega)*} E^{(\omega)} = (2\pi\hbar\omega/V) n \quad \dots(11)$$

$$\tan^{-1} (E_i^{(\omega)}/E_R^{(\omega)}) = \phi. \quad \dots (12)$$

If the fields are quantized, the coefficients  $E_{m\mu}^{(\omega)}$  are no longer c-numbers. In such a case, one finds the familiar uncertainty relation between the photon number  $n$  and phase  $\phi$ :

$$\Delta n \Delta\phi \geq 1. \quad \dots(13)$$

In the representation in which states are eigenstates of the photon number operator  $n = (V/2\pi\hbar\omega) E^{(\omega)*} E^{(\omega)}$ , one has the explicit operator properties

$$\begin{aligned} E^{(\omega)} |n\rangle &= (2\pi\hbar\omega/V)^{1/2} \sqrt{n} |n-1\rangle; \\ E^{(\omega)} |n\rangle &= (2\pi\hbar\omega/V)^{1/2} \sqrt{n+1} |n+1\rangle. \end{aligned} \quad \dots(14)$$

Light prepared in any coherent state  $|\alpha\rangle$ , which is the eigenstate of the destruction operator  $(V/2\pi\hbar\omega)^{1/2} E^{(\omega)}$ , with eigenvalue  $\alpha$ , is said to be in the minimum uncertainty state, since in such a case

$$\Delta n_c = \Delta\phi_c, \quad \Delta n_c \Delta\phi_c = 1. \quad \dots(15)$$

The state of light photons can also be prepared in the so called squeezed quantum state<sup>6</sup> in which the uncertainty in either  $\Delta n$  or  $\Delta\phi$  (or uncertainty in any other combination of these conjugate variables) is less than the minimum uncertainty  $\Delta n_c = \Delta\phi_c$  of the coherent state. Because of this property, such states have very special significance in devising extremely sensitive optical detectors. Obviously, the uncertainty in the corresponding conjugate linear combination will be higher than in the minimum uncertainty state.

In general,  $P$  is not a linear function of  $E$ . To describe nonlinear optical processes<sup>7</sup> in a material medium, the electronic polarization (which is the dominant part at optical frequencies) can be expanded in perturbation series, written symbolically as

$$P = P^{(1)} + P^{(2)} + P^{(3)} + P^{(4)} + \dots \quad \dots(16)$$

$$P_i^{(1)}(\omega) = \sum_{j=1}^3 \chi_{ij}^{(1)}(\omega) E_j(\omega);$$

$$P_i^{(2)}(\omega_1 + \omega_2) = \sum_j \sum_k \chi_{ijk}^{(2)}(\omega_1, \omega_2) E_j(\omega_1) E_k(\omega_2) \quad \dots(17)$$

and

$$P_i^{(3)}(\omega_1 + \omega_2 + \omega_3) = \sum_j \sum_k \sum_l \chi_{ijkl}^{(3)}(\omega_1, \omega_2, \omega_3) E_j(\omega_1) E_k(\omega_2) E_l(\omega_3), \text{ etc.} \quad \dots(18)$$

In the electric dipole approximation being used here, this expansion is strictly valid only for nonresonant passive systems in which fields  $E$  are weak enough so that the relevant Rabi frequencies satisfy the relation

$$|(\mu_{ng} \cdot E)/\hbar| \ll |\omega - \omega_{ng} - i\Gamma_{ng}| \quad \dots(19)$$

where  $\mu_{ng}$  are the electric dipole transition matrix elements between relevant states  $n$  and  $g$ ;  $\hbar\omega_{ng}$  and  $\Gamma_{ng}$  are the corresponding energy difference and damping rate for the transition. In such a situation, the above expansion is an expansion in powers of  $|E/E_a|$  where  $E_a \sim 10^6$  esu is of the order of atomic fields. This implies that  $x^{(2)} \sim 10^{-6}$  to  $10^{-7}$  esu,  $x^{(3)} \sim 10^{-12}$  to  $10^{-14}$  esu, etc. (since  $x^{(1)}$  is of the order of 1), unless one is close to the resonance condition i.e., close to the vanishing of the right hand of Eq. (19) with respect to some relevant intermediate states. The possible nonvanishing components of  $\chi_{j_1 j_2 \dots j_n+1}^{(n)}$  and their mutual relationships depend strongly on the symmetry properties of the material. For example, for centro-symmetric systems,  $x^{(2)}, x^{(4)}$  etc., are identically zero in the dominant electric-dipole approximation being considered here. For three-wave interactions involving  $x^{(2)}(\pm\omega_1 \pm \omega_2)$ , one can generate fields at frequencies  $\omega_3 = |\pm\omega_1 \pm \omega_2|$ , i.e., the sum frequency, the difference frequency, the second-harmonic, and the rectified field, only in noncentro-symmetric systems like lithium niobate, GaAs, etc. In the absence of appreciable absorption, the transfer of energy can be quite efficient if there is the so-called phase matching, i.e., if propagation vectors of the interacting waves are arranged such that  $q_1 + q_2 = q_3$ , e.g., by using suitable bi-axial crystals. In such a case, one can manage to get parametric amplification of signals with very little noise. Novel sensitive parametric detectors involving squeezed states are based on the 3-

wave parametric interaction. For the case of 4-wave interaction described by  $\chi^{(3)}(\pm \omega_1 \pm \omega_2 \pm \omega_3)$ , one can use both centro-symmetric systems and non-centro-symmetric solids. These processes include third-harmonic generation ( $\omega + \omega + \omega$ ), intensity-dependent refractive index ( $\omega - \omega \pm \omega$ ), frequency mixing ( $\omega_1 + \omega_1 - \omega_2$ ), etc. Because of the intensity dependent-refractive index, one has the possibility of self-focussing of laser beams passing through such a medium. If such a nonlinear medium with intensity dependent refractive index is kept in between two mirrors (Fabry-Perot etalon), one obtains a bistable switching device for transmitted intensity when the incident beam intensity is varied to sweep across a cavity resonance frequency.

Also, using the fourwave interaction process involving the induced polarization  $\chi^{(3)}(-\omega, +\omega, +\omega): E_1^*(\omega) E_2(\omega) E_3(\omega)$  in which the pump beams 2 and 3 counter propagate with  $q_2 = -q_3$ , one can generate the field 4 at frequency  $\omega$  which is exactly phase-conjugate to the incident signal field  $E_1(\omega)$ . This leads to the fabrication of phase-conjugated mirrors<sup>5,9</sup> for removing distortion of images.

If some of the interacting fields are not at optical frequencies, one gets other familiar nonlinear effects. For example, for  $\omega_s \ll \omega$  (optical),  $\chi^{(2)}(\omega, \omega_s): E(\omega) E(\omega_s)$  is nothing but the linear electro-optic effect (Pockel's effect) which is used for modulating optical waves. The coefficient  $\chi^{(3)}(\omega, \omega_s, \omega_s): E(\omega) E(\omega_s) E(\omega_s)$  relate to the Kerr effect and  $\alpha(\omega, 0): E(\omega) B(0)$  to the magneto-optic (Faraday) effect. The magneto-optic effect allows an external magnetic field to control the plane of polarization of the optical wave, which is useful to fabricate "isolators" in propagating optical circuit. Other types of nonlinearity involve modulation of  $\chi^{(1)}$  by optical phonons or acoustic waves, with amplitudes  $Q_R$  and  $u$  respectively:

$$P^{NL} = \left( \frac{\partial \chi^{(1)}}{\partial Q_R} \right) : E(\omega) Q_R(\omega_s) + \left( \frac{\partial \chi^{(1)}}{\partial u} \right) : E(\omega) u(\omega_s) \quad \dots(20)$$

Here, the first term is the Raman nonlinearity and the second term is the acousto-optic nonlinearity of the Raman-Nath theory. These effects are used, respectively, to fabricate Raman lasers at shifted frequencies and to modulate an optical wave by acoustic signal directly.

## OPTOELECTRONICS: PRESENT SCOPE

In information processing technology involving communication, computing, data processing and storage, optics has obvious advantages, compared to the conventional electronics, because of the possibility of large-bandwidth, inherent two-dimensional processing, ease in its propagation in space or in weakly absorbing materials and very weak interaction between waves at different frequencies (no "cross-talk"). The conventional field of micro-electronics based on the control and manipulation of carriers in semiconductors, with sequential processing, is at present still the workhorse of the information industry. But, the search for faster and compact information processing systems is a continuing one. In this connection, it should be noted that conventional steady state semiconductor devices are limited in their speed by the carrier relaxation time of about  $10^{-9}$  sec. One can increase the speed to  $10^{12}$ /sec (picosecond response level) using novel semiconductor devices based on quantum-well structures, high mobility transistors or resonant tunneling devices, and one can also miniaturize and increase circuit density on a single semiconductor chip to meet the increasing demands of large memory and faster communication between different parts of the chip, but it cannot proceed beyond a certain limit because of the severe bottleneck due to slow removal of dissipated heat in such small structures. The basic aim of new technology involving the light wave is to fabricate photonic devices in which optical photons instead of electronic carriers are manipulated and controlled for the desired task. However, at present, the information technology is evolving towards a hybrid technology in which the best of both microelectronics and photonics are to be used for performing specific functions in the complete system. This technology can be described very aptly by the word "optoelectronics". As of now, the light modulation speeds are not fast enough and the optical detectors are not sensitive enough to take the full advantage of large optical band width and other useful properties of light. We are still far away from the dream technology of "photonics" in which except, possibly for only the input/output devices, all other devices will be purely photonic without too many conversions from electrons to photons and reconversions from optical photons to electrons.

The major applications of opto-electronics are related to (i) information storage, (ii) optical signal and image processing, (iii) optical communications, (iv) computing, (v) metrology and control, (vi) industrial, medical and defense related systems, and (vii) basic sciences. Here, because



of obvious limitations, we will not be able to deal with any of these applications in detail. However, in what follows, we propose to present a bird's eye-view of the progress made in this direction on the first five topics listed above.

### *Information Storage*

In any storage device, one wants to record information at very high speed, with proper indexing, and retrieve the desired part whenever needed, again at very high speed. Whereas the old storage methods involving book prints and photographic films are too slow for present-day needs, the modern electronic storage based on magnetic memory and semiconductor devices may not be large and fast enough for the next decade. Optical recording and reading technology is developing rapidly to meet this demand. The simplest device in this context is the optical compact disk (CD). This can be in the form of "read only" mode or in the WORM (write once, read many times) mode for archival storage or in the form of erasable optical disks which can be used to write and read many times. Digital papers with storage capacity of 5 Megabits(Mb) per side per channel are now available commercially, and erasable 130mm optical disks with 550 Mb per side, having response time of approximately  $10^{-9}$  s, are at the final stage of development. The optical storage technology, in general, is based on one of the familiar optical processes involving electro-optic, magneto-optic, bump-pit formations on disks, bistability, holography, thermomagnetic, or localized electron trapping phenomena in suitable materials. Holographic recording without digital conversion can be used directly for the optical memory of a computer. It is being said that erasable (rewritable) magneto-optical disks (MOD) will finally replace the present-day magnetic recording, since the optical recording is a non-contact process. As of now, the second-generation two-head three-beam magneto-optical mass-storage systems have rewritable optical-disk drive which can store  $10^{12}$  bits of information (~5 million newspaper pages!) with recording rate of  $2 \times 10^6$  bits per second.

### *Signal and Image Processing*

In conventional electronics, signal and image processing are done *via* digital electronics, implying conversion of the image or the signal into a digital form for processing and subsequent reconversion into the original mode. Unfortunately, for many applications present-day electronic digital processing is not fast enough. The direct optical processing has the obvious advantage in this respect because it can be done in two-dimensional space. Even if the speed for individual processing in optical and electronic

systems, remains the same, although in future this need not be so, optical processing is inherently faster. A certain amount of parallelism (parallel electronic processing) can, of course, be introduced in the conventional technology, but eventually it would be impossible to compete with optical processing. One can use either a spherical lens which performs two-dimensional F T, or stacks of cylindrical lenses each of which performs one dimensional F T. Using a set of lenses and beam splitters, and special masks and filters corresponding to particular spatial frequencies in the image, one can perform most of the logical functions like, addition, subtraction, multiplication, division, averaging and differentiation of the source functions, needed in image processing. The processing can be done by either using coherent light or less demanding incoherent light. By comparing input pattern with a given reference pattern, this technique is used to fabricate low-speed hybrid processors for computers. Other important applications of optical image processing are in spatial spectrum analysis, matched filtering, radar, pattern and character recognition, computer-aided vision, image enhancement, aerial photography, metrology, and missile guidance.

### *Optical Communication*

The idea of modern optical communication system<sup>11</sup> is not new. The greatest handicap in its implementation has been the difficulty in modulating the optical wave at a very high speed to take advantage of its large bandwidth. Of course, from the days of "light-houses" involving very low speed of modulation, we have come a long way, with the present-day maximum modulation speeds of  $10^{11}$  to  $10^{12}$  per sec. But, we have still to go quite far in increasing the modulation and detection (demodulation) speeds to take full advantage of the large band-width. As in any such technology, the complete optical communication system consists of a transmitter which transmits the light wave modulated by the signal, the transmission medium in which the modulated wave propagates, and a receiver where the wave is detected and demodulated to isolate the signal. The transmitter end consists of an optical source which is modulated by the signal by a suitable technique and an optical amplifier which amplifies it before transmission. The transmission medium for optical wave is either low-loss optical fibres for terrestrial communications or free space (with no atmospheric absorption and distortion) for space communications. For transmission of light through very large distances, one also requires repeaters consisting of a receiver and a transmitter. The receiver end detects and demodulates the input wave, and after the process of pulse shaping and amplification to boost the signal

quality the signal is transmitted through the transmitter end of the repeater. With the rapid progress being made in fabricating low-loss multi-mode doped-silica glass fibres, the loss near  $1.55\ \mu$  wavelength has come down to about 0.2 dB per kilometer (km) at  $10^{12}$  Hz bandwidth. This is much better than attenuation levels of 10 dB per km at 10 MHz bandwidth for coaxial cables and about 1 dB/km at  $10^{10}$  Hz bandwidth of waveguide transmission. Thus, the required distance between two repeaters is very much larger for the case of optical-fibre transmission than in the case of conventional coaxial cable or microwave transmission. It is because of this great cost-saving in transmission that has prompted the long-distance communication industry to shift to optical transmission although they continue to use conventional electronic transmitters and receivers at low-bandwidth levels. Even though one has to introduce additional devices to modulate LED (light-emitting diode) source by the low-frequency wave containing the signal at the transmitter end and a PIN photo-diode detector at the receiver end, optical-fibre transmission turns out to be much cheaper. The ideal "optical" communication system should, of course, have optical transmitters and optical receivers with modulation and demodulation speeds of greater than  $10^{12}$  to  $10^{13}$  Hz. All attempts in this field are towards the development of this challenging frontier technology.

As emphasized earlier, the level of progress in large scale applications of optical technology to the information processing industry is linked to the development of suitable compact optical sources, which can be modulated at very high speeds, and sensitive optical detectors with fast response time and high sensitivity in the spectral range of interest. Because of the requirements of small size and easy integratability with other optical and electronic devices and elements on a chip, the development of suitable semiconductor diode lasers and sensitive photo-diode detectors is the major task being pursued at present. For fibre-optic links, which works best in the region around the wavelength of  $1.3\ \mu$  to  $1.55\ \mu$ , as far as their attenuation and frequency dispersion are concerned, proper sources and detectors have to work in that infrared range. However, it must be pointed out immediately that for many other important commercial applications, such as laser printers, optical memories, compact disk players, industrial sensors, high-definition television and colour displays (red, green and blue), novel semiconductors and other solid-state lasers in the visible range are also in great demand.

In optical communication systems, one can use either incoherent light sources like LED's (light emitting  $p$ - $n$  diodes) based on InGaAs/InP, or AlGaAs/GaAs coherent  $p$ - $n$  diode lasers. However, to avoid the corresponding distortions due to mode dispersion as well as residual material dispersion in the optical fibre, a single-mode laser with narrow spectral width, coupled to a single-mode fibre<sup>12</sup>, is the best solution for advanced systems in which modulation speeds exceed  $10^9$  bits per second. Several types of dynamic single-mode lasers based on distributed feed-back or distributed Bragg reflection have been fabricated for this purpose. A direct modulation of light emitted in LED is possible by modulating the input current. However, because of large carrier lifetimes ( $\tau \simeq 10^{-9}$ s), the modulation bandwidth is usually limited to  $\Delta f \simeq 1/2\pi\tau \leq 10^9$  Hz. Higher modulation speeds of  $10^9$  Hz to  $10^{11}$  Hz have been achieved in the case of the dynamic single-mode laser diodes. New coherent optical sources and amplifiers based on novel quantum-well and superlattice semiconductor structures are being developed to push modulation speeds beyond  $10^{11}$  Hz. Modern optical detectors which can be easily integrated with semiconductor technology are based on the reverse-biased InGaAs/InP PIN photo-diodes (opposite of laser-diodes, photon to electron converter instead of electron to photon converter). However, in such photo-diodes there is no gain as in conventional sensitive photo-multiplier tubes (PMT) where multiplication of primary electrons is accomplished by a series of dynodes without introducing additional noise. A considerable amount of gain can be introduced in the so-called avalanche photo-diodes (APD) but it is quite noisy compared to PMT. What is required is a compact solid-state version of PMT. New ideas based on band-gap engineering<sup>13</sup> are being tried out to develop a solid state PMT, with response time  $10^{-12}$  s or less. Finally, having developed all the necessary discrete elements<sup>14</sup> (laser sources, modulators, switches, couplers, filters, dividers, connectors lenses, optical dielectric waveguides, isolators, amplifiers<sup>15</sup>, detectors, demodulators, etc.), they have to be integrated on a chip. However, opto-electronic integrated circuit (OEIC) technology<sup>16</sup> at present is still in infancy. One is able to put only a small number of channels ( $4 \times 4$  or  $8 \times 8$ ) on a single chip in integrated coherent optical receivers and transmitters of to-day. There is a long way to go to catch up with the very large scale integrated circuits of conventional semiconductor electronics.

### *Optical Computing*

The idea of optical computing, with two-dimensional parallel processing, is extremely fascinating. In modern electronic computing systems, a certain amount of parallel processing is being introduced to overcome the usual von Neumann bottleneck due to serial accessing to data in the memory and sequential processing. However, for faster and large computing power, optical computing is the ultimate answer. Unfortunately, most optical processors available to-day are analog hybrid processors which cannot take full advantage of the parallelism of optical computing; the digital computation being actually performed electronically. Any computer consists of an input device, an output device a central unit containing the central processing unit (CPU) and memory and interconnections between them. As far as interconnects are concerned, optics is far better. They are being used already for internal communications between different units of a computer, in new main-frame systems. The development of optical storage and memory is also satisfactory. However, the progress made in fabricating suitable optical logic gates, etc., for an acceptable optical digital processing unit is quite slow<sup>18</sup>. A complete optical computer is still a dream<sup>19,21</sup>. But, the days of hybrid computers in which main processing unit is still electronic but memory and innerconnects are optical, are not too far.

### *Metrology and Control*

Light has always been used as sensor in measurement and control systems. With the rapid progress made in optical technology during the last thirty years, optoelectronics is playing a major role in geometrical control systems and physical parameter sensors. For visual systems in which information regarding the position, orientation, shape and size of the object is sought for recognition, inspection and manipulation, simple imaging sensors<sup>22,23</sup> like CCD cameras and various position-sensitive detectors are used. Optical fibre sensors<sup>24</sup> for determining various physical parameters like temperature, pressure, position, the presence of chemical products, magnetic and acoustic fields, etc., are already available commercially.

One can go on and on, narrating various other important applications of optoelectronics, particularly its applications in the medical field (ophthalmology, microsurgery) and in not so peaceful, the military field (laser-guided bombs and missiles, night vision devices, lidars). However, instead of doing that, in what follows we will briefly examine the future prospects of graduating from the technology of optoelectronics to the technology of photonics.

## MARCH TOWARDS PHOTONICS

In the preceding section, we have been describing the development of the new hybrid technology of optoelectronics. It is quite certain that in the next five to ten years, this technology will advance to the extent that one will be able to take full advantage of the large bandwidth available in optics. Novel devices based on nonequilibrium processes<sup>25</sup> and band-gap engineering<sup>13</sup> are expected to become available which will allow modulation of compact optical sources at  $10^{13}$  Hz or so and detect optical photons with response time less than  $10^{-12}$  -  $10^{-13}$  s. In future, integrated opto-electronic circuits will also be the order of the day, although large scale integration (LSI) on the scale possible at present in semiconductor microelectronics technology will take considerably more time. The same thing is true for the development of fast optical processors for digital computational logic. However, the fundamental point remains to be tackled! If we are going to use optical photons to do everything what electrons do in the conventional electronics technology, why not develop complete photonic systems, without converting and reconverting electrons to photons and photons to electrons in the hybrid technology of optoelectronics? Except for the input and output devices, photons should take over all the tasks. That will be the real technology of "photonics"!

The march towards real "photonics" has already begun. The questions<sup>26-28</sup> regarding overcoming the signal shot-noise limitations and detection of each bit of information by 1 photon only, quantum mechanical computers, fundamental limitations of computing, and direct modulation and demodulation of light are already being answered satisfactorily at least at the laboratory level. In a quantum nondemolition (QND) scheme, one can detect photons via the 3-wave parametric interaction or similar other nonlinear optical processes in which no additional noise is added in the process of detection. The use of optical squeezed states<sup>29,30</sup> to increase detection efficiency beyond the quantum limit and light propagation in the form of optical solitons in optical fibres and detecting them via noiseless parametric interaction with a probe optical soliton, are bringing in new ideas for the future photonic technology. In this direction, the next decade should turn out to be the most exciting and challenging period for optics, both from the point of view of basic science as well as from the point of view of the technology at the frontiers of knowledge.

## ACKNOWLEDGEMENT

It is a great pleasure to thank the Indian National Science Academy for asking me to deliver the Jawaharlal Nehru Centenary Lecture for the year 1991. I am also thankful to the Punjab University, Chandigarh, for arranging the talk at Chandigarh. I feel greatly honoured to get this opportunity to pay my homage and humble tribute to one of the greatest builders of the Indian democracy, who did so much for science in India, and who was also one of the finest of human beings. I very much hope that we can build India of his dream through science and technology.

## References

- 1 K Iizuka *Engineering Optics, Springer Series in Optical Sciences, Vol 35, Second Edn Springer-Verlag Berlin* (1987)
- 2 S S Jha in *Resonances* (Ed M D Levenson *et al*) World Scientific Singapore (1990) 259; A S Vengurlekar and S S Jha *Phys Rev B* **41** (1990) 1286
3. *Quantum Optics, Experimental Gravitation and Measurement Theory* (Ed. P. Meystre and M O Scully) Plenum, New York (1981)
4. Z Y Ou and L Mandel *Phys Rev Lett* **61** (1988) 50
5. A Yariv *Optical Electronics* (3rd Ed) B Holt, Rinehart and Winston, New York (1985).
- 6 R E Slusher, L W Hollberg, B Yurke, J C Mertz and J F Valley *Phys Rev Lett* **55** (1985) 625
7. Y R Shen *The Principles of Nonlinear Optics*, Wiley-Interscience, New York (1984)
8. A Szoke, V Daneu, J Goldhar and N A Kurnit *Appl Phys Lett* **15** (1969) 581
9. A Yariv and D M Pepper *Opt Lett* **1** (1977) 16
- 10 L K Van Geest In: *Photoelectronic Imaging Devices* (Ed B L Morgan) Academic, New York (1988) 1
- 11 J Gower *Optical Communication Systems* Prentice Hall, Englewood Cliffs, NJ (1984); A K Ghatak and K Thyagarajan *Optical Electronics* Cambridge Univ. Press (1989)
- 12 L B Jeunhomme *Single-Model Fiber Optics* Marcel Dekker New York (1990)
- 13 F Capasso *Science* **235** (1987) 172; Also, see *Picosecond Optoelectronic Devices* (Ed. C H Lee) Academic Orlando Florida (1984)
- 14 T Uchida *Proc 20th European Microwave Conference Microwave Exhib & Publ., Tunbridge Wells, U K, Vol. 1* (1990) 120

- 15 B J Ainslie *J Lightwave Technol* 9 (1991) 220
- 16 R M Ash, R Goddellow and A C Carter *Proc SPIE - Int Soc opt Engng (USA)* 1043 (1989) 214
- 17 D Docoster and J P Vilcot *Proc 19th European Solid State Devices Res Conf* Springer Verlag Berlin (1989) 329
- 18 S D Smith *Appl Opt* 25 (1986) 1550
- 19 W T Cathey *Proc 9th int Electronics Packaging Conf* Int Electr & Packaging Soc San Diego CA (1989) 1051
- 20 S Hunter, F Kiamilev, S Esener, D A Parthenopoulous and P M Rentzepis *Appl Opt (USA)* 29 (1990) 2058
- 21 A L Mikaelyan *Sov J Commun Tech Electron (USA)* 35 (1990) 21
- 22 J A Tyson *J Opt Soc Am A* 7 (1990) 1231
- 23 R M Measures *Laser Remote Sensing* John Wiley & Sons New York (1984)
- 24 D A Krohn *Fiber Optics Sensors - Fundamentals and Applications* Instrument Soc America (1988)
- 25 A S Vengurlekar and S S Jha *Phys Rev B* 43 (1991) 12454
- 26 R P Feynman *Found Phys* 16 (1986) 507
- 27 C H Bennet and R Landauer *Scient Am* 253 (1985) 48
- 28 H J Caulfield and J Shamir *J Opt Soc Am A* 7 (1990) 1314; *Appl Opt* 28 (1989) 2184
- 29 A Piskarskas, A Stabins and A Yankauskas *Opt Spectrosc (USA)* 66 (1989) 434
- 30 S A Akhmanov, A C Belinski and A S Chirkin, *Opt Spectros (USA)* 66 (1989) 432





**Maroli Krishnayya Chandrashekaran** (b. 4 January 1937) did Ph.D.(1964) from University of Madras and D.Sc.(1985) from Madurai Kamaraj University. He is Professor at School of Biological Sciences, Madurai Kamaraj University, Madurai (1980- ).

Chandrashekaran's contributions relate to: rediscovery in India of tidal and circadian rhythms in a crab; elucidation of light relations of circadian rhythm in *Drosophila*, first reports of social synchronization of circadian rhythm in a bat; and maternal entrainment of circadian rhythm in field mouse. His current research concerns circadian

rhythms, time sense and menstrual cycles in humans under prolonged social and temporal isolation. He has authored 'Biological Rhythms' (Bharatiya Vidya Bhavan, 1985); 'Biological Rhythms' (Tamil edn) 'Uyiriyal Talamurai' (1991); co-authored 'Animal Behaviour: Basic Experiments in Neurophysiology' (Macmillan India, 1987).

Chandrashekaran is Member of Guha Research Conference; Member, Indian Academy of Neuroscience, Member, Society of Light Treatment and Biological Rhythms (USA); Fellow, Indian Academy of Sciences, and National Academy of Sciences (India). He is the recipient of Shanti Swarup Bhatnagar Prize (1979); UGC National Lecturer (1983-84); Sir J.C. Bose Award (UGC-Hari Om Ashram Trust) (1989), UGC National Fellow (1984-86); Jawaharlal Nehru Birth Centenary Lecture Award (INSA) (1991); SERC Lecturer (DST) (1992).

*Maroli Krishnayya Chandrashekaran was elected to the fellowship of the Academy in 1988.*

**Address:** Professor and Head, Department of Animal Behaviour and Physiology, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021, Tel: (0452) 85216; Telex: 445-308 MKU-IN; Fax: 91-452-85205; Grams. University, Madurai; Res: 1/64, Faculty Quarters, Madurai Kamaraj University, Madurai 625 021; Tel: (0452) 85447.

## **SOCIAL CUES AND CIRCADIAN RHYTHMS IN BATS, MICE AND HUMANS**

M K CHANDRASHEKARAN

It was known possibly, ever since the French astronomer demonstrated the endogenous nature of diurnal or circadian rhythms in a plant in 1729, that the light/darkness (LD) cycles entrained or regulated them <sup>(1)</sup> Experimental work of over a 100 years has established LD cycles to be ubiquitous entraining agents or zeitgebers of circadian rhythms. The identification of the two is so close that chronobiologists were also often expected to be members of the societies of photobiologists and photochemists.<sup>(2)</sup>

### **DIFFERENT KINDS OF ZEITGEBERS**

Zeitgebers may be strong as in the case of LD cycles, weak as in food: starvation (or restricted feeding cycles) or subtle as when cosmic factors, electrostatic fields and magnetostatic fields are invoked. Biological clocks are obviously versatile and opportunistic and may derive their time cues from whatever source they can. It must, however be pointed out that evidence for the efficacy of 'subtle' zeitgebers is tenuous and equivocal and has been centre of interest of only one laboratory in the past and possibly none at present.

### **SOCIAL ZEITGEBERS**

The role of social interaction in synchronizing activity rest patterns of several species of birds and mammals seems only too clear from the great synchrony with which flocks of animals leave their roosting sites to forage at sunset <sup>(3)</sup>. Many colonies of bats seem to emerge from their caves as if by coup and the colony emergence presents the appearance of billows of smoke. The individual members then scatter and fly apart even if they do so in the same direction and may eventually forage in the same general area.

Halberg and co-workers (1954)<sup>(4)</sup> worked with blinded mice *Mus musculus* and showed that they *entrained* to LD cycles *only* if normal mice were housed in the same room. The authors postulated that auditory and olfactory factors mediated social synchronization in this mice. It was

reported that female house sparrows entrained their perch-hopping rhythms to the male songs played back to them through a tape-recorder. They treated the song part of the 24h cycle as day and silence as night in experiments which were performed in constant light (LL) for convenience.

One of the earliest reports to impute social synchronization among conspecifics was that for the mice of the genus *Peromyscus*. Subsequent reports described similar effects among male chevrotian antelopes, wolf-coyote hybrids, beaver colonies of *Castor canadensis* macaque monkeys, sexual cyclicity of female mammals, and so on.<sup>(5)</sup>

The experimental tradition in social chronobiology is of recent origin. The contributions of original findings on the social synchronization of circadian rhythms in bats and<sup>(6-10)</sup>, mice<sup>(11-13)</sup> and humans<sup>(14,15)</sup> from our laboratory in Madurai are landmarks in the literature on the subject.

## BATS

Field ethological studies on the behaviour and biological clocks of microchiropteran bats have been carried out in Madurai (9°58' N lat 78° 10' E long) inside natural caves

A colony of about 400 to 500 insectivorous bats of the species *Hipposideros speoris* inhabit a 'true cave' (that is, a cave with just one opening) close to the Madurai Kamaraj University campus. The cave has several labyrinthine ramifications 15 to 50 m deep, that is, from its mouth. The bats use several of these pockets as their daytime roosting place. Most roosting spots show great constancy of temperature (27°C) and humidity (95%) and the darkness is absolute, day and night. Within the cave, there is hardly any clue to the passing of time. Yet foraging flight occurs regularly 10 to 15 minutes after sunset. It is clear that the animals are aware of the sunset.

The sequence of events culminating in such regular outflights are as follows. The bats awaken well before sunset. They then stretch, preen themselves and undertake short flights within the dark recesses. They fly into an outer chamber that opens to the world outside through a small window-like mouth; in this chamber they "sample-light". When it gets sufficiently dark outside they fly out.

The first question that arose was: should each bat sample light for itself or would conspecifics relay the information to those in the interior

region? The only way to find this out was to keep a few bats prisoners in their own cave 40 m deep and record their flight activity. Flight cages with writing stylets and mechanically wound thermohygrograph drums were used in the study. The activity data of the experiment which lasted 50 days show that even the captive bats recognized the time (Fig. 1). Evening after evening the captive conspecifics inside the cage flew about. Interestingly they also responded to the several bat returns. There can be little doubt that the captive bats are responding to social cues. That means the free-flying conspecifics apparently transmit the news of the sunset.

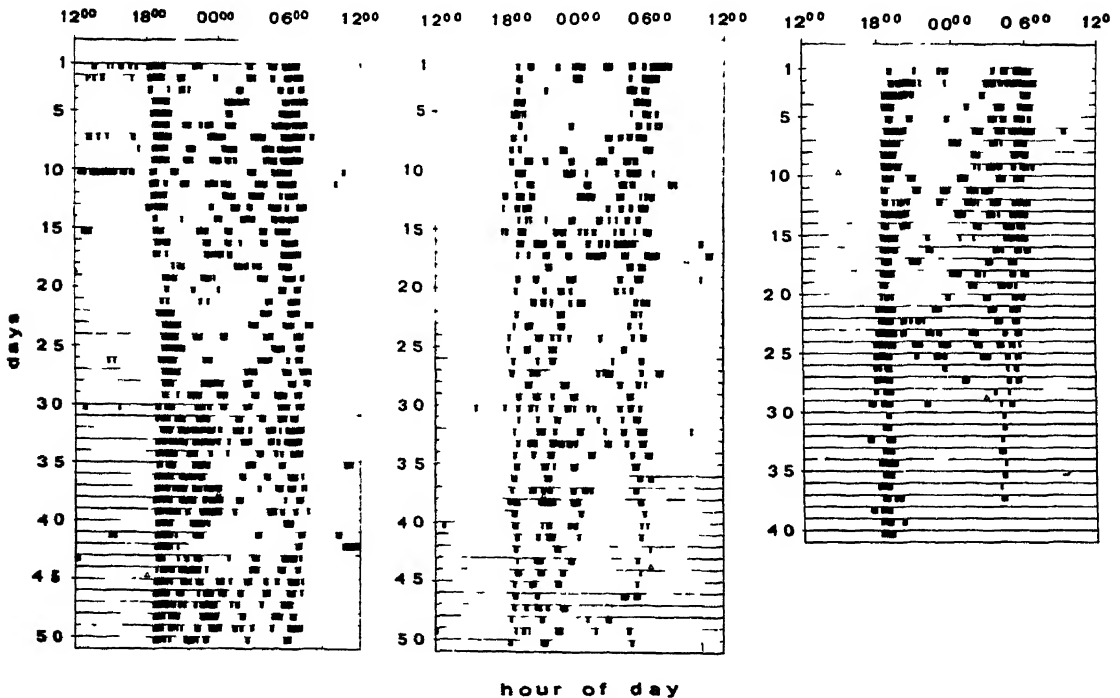


FIG 1 The social entrainment of three bats (*H. speoris*) held captive inside a cave under natural darkness in the presence of free flying conspecifics. The vertical bands and horizontal lines represent the activity and rest periods, respectively. The data are schematized from the original recording and presented one below the other for successive days. (After ref. 6).

A second experiment further probed the hypothesis of social synchronization of biological rhythms in this bat. If freeflying members of the bat colony transmitted the message of time, should not a solitary bat in a solitary cave be helpless? This experiment was carried out in a different cave where there were virtually no bats living. (The word virtually was

used since 3 bats of the species *Hipposideros speoris* did live there and these 3 animals had to be exterminated since it is nearly impossible to make these animals abandon their day time roosts since they would always return on the rebound with the aid of their "spatial memory"). A solitary male *Hipposideros speoris* bat was held in this cave devoid of any other bat and his activity was recorded over a period of 50 days. The solitary bat is indeed helpless as regards time (Fig. 2). Its rhythm "freeruns" and gains about 20 minutes every day.

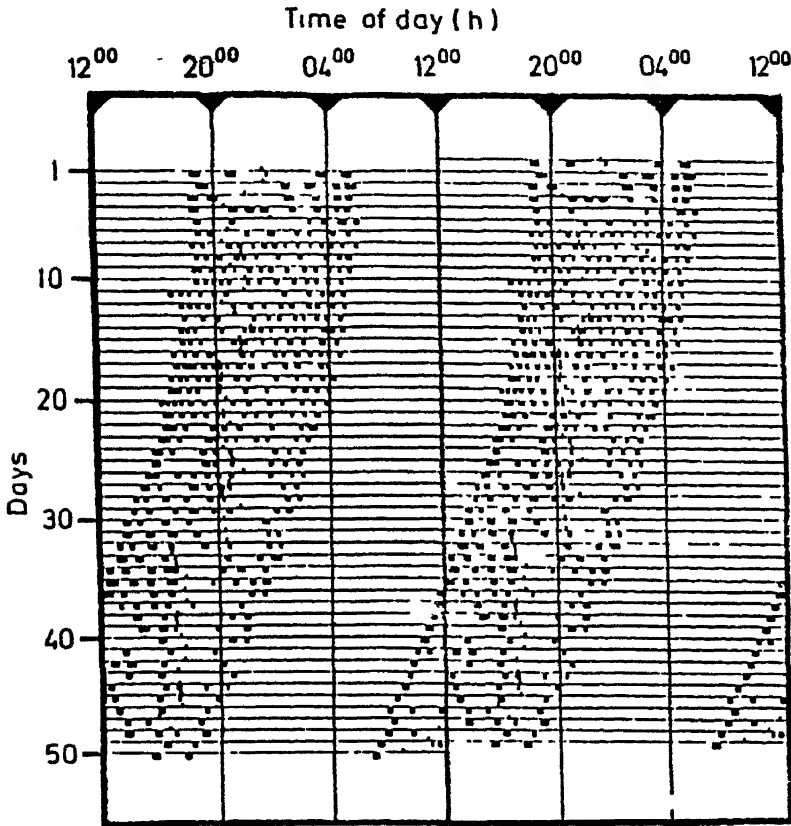


FIG 2 A double plot of the free running rhythm of the flight activity of a solitary male *H. speoris*. The recording was done inside a cave without conspecifics. Other details as in Fig. 1. (After ref. 6).

The following factors might mediate in the synchronizing process:

- (1) flight noise of free-flying members
- (2) some unknown pheromonal or chemical message emitted by the flying bats and
- (3) ultrasound that these bats emit.

An interesting feature of this experiment must be pointed out. The solitary cave without any bats happened to be close to a water-hole where crows and mynas flocked during day time to drink. The investigators could hear them during the day and the stridulation of crickets and the

croaking of frogs during the night. To that extent the cave was not entirely without social time cues, only the cues came from other species of animals.

The finding that non-specific social cues did not entrain the circadian rhythmicity in *Hipposideros speoris* egged us on to perform another experiment inside the *Hipposideros speoris* cave. In this experiment we held a non-hipposiderid bat, another insectivorous microchiropteran, *Taphozous nudiventris kachhensis*, captive 40m deep in the cave and measured his daily flight activity. It turned out very surprisingly that the rhythm of the alien individual bat, free ran (Fig. 3). The captive and single *Taphozous* bat rhythm was not entrained by the bustle and wing-beats of some 400-500 *Hipposideros speoris* bats which flew out evening after evening around 7 p.m. The possibility that even the wingbeat noise of these 2 species of bats might vary cannot be overlooked; even pheromones, if any involved, are known to be species specific. As regards the ultrasonic components *Taphozous nudiventris kachhensis* emits in the region of 80 kHz (and *H. speoris* emits in the region of 135 kHz).

The first two factors are present continuously inside the caves; however, they could steeply intensify. While not ruling this out completely, we find the ultrasonics factor as the most interesting zeitgeber.

*H. speoris* is silent to humans, whereas *Taphozous* can emit audible vocalizations also. Ultrasound has been demonstrably used by bats only to catch prey, avoid obstacles and in such other immediate behavioural contexts. If we can establish the zeitgeber role of ultrasonics in bats we would be giving it a new and social dimension. Experiments to unravel the phenomenon (with tape-recorded ultrasound sound-proof experimental cubicles, etc.) are underway in Madurai.

In other series of experiments the bats were exposed to a conflicting zeitgeber situation. This, however, may not ever occur in nature. The captive bats inside the cave were held in continuous light of 10-20 lux. We noticed that the circadian rhythm in the flight activity of three male *Hipposideros speoris* bats freeran with a period of >24 h. The freeruns indicate that whatever social cues that prevailed inside the cave are not reaching the clock or in fact that LL conditions might have indeed actively abolished the social synchronization.

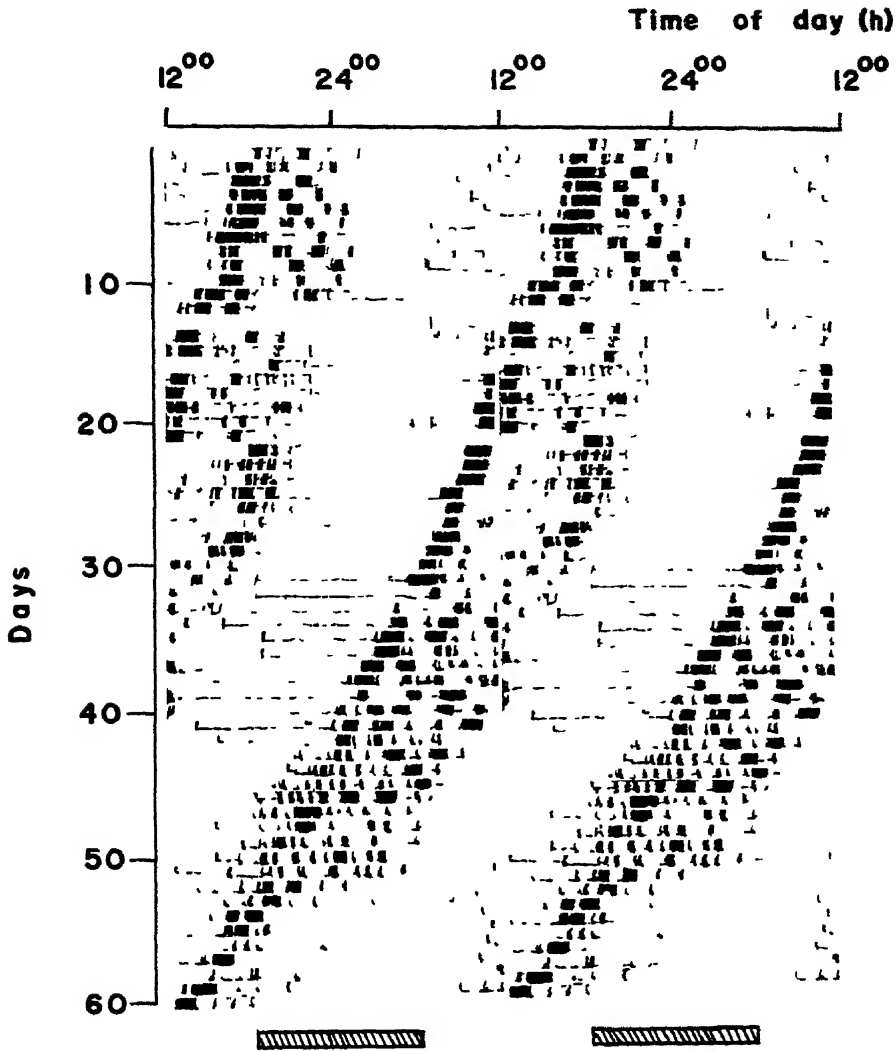


FIG 3 The free running rhythm of a captive *T.n. kachhensis* recorded inside a cave in the presence of about 500 individuals of a colony of free flying *H. speoris*. The data shows the original felt pen tracings. The hatched area at the bottom indicates the duration in which the hipposiderid colony would be active, flying out of the cave during dusk and returning during dawn hours. (After ref. 9).

### MICE

We have also reported for the first time that the mother field mouse *Mus booduga* can behaviourally entrain the circadian rhythms in the running activity of her pups. Cycles of 12 h presence and 12 h absence of mother *Mus booduga* entrain the circadian rhythm in the locomotor activity of her pups such that the pups rest in her presence and are active in her absence (Fig. 4). We wanted to determine whether this maternal entrainment arises

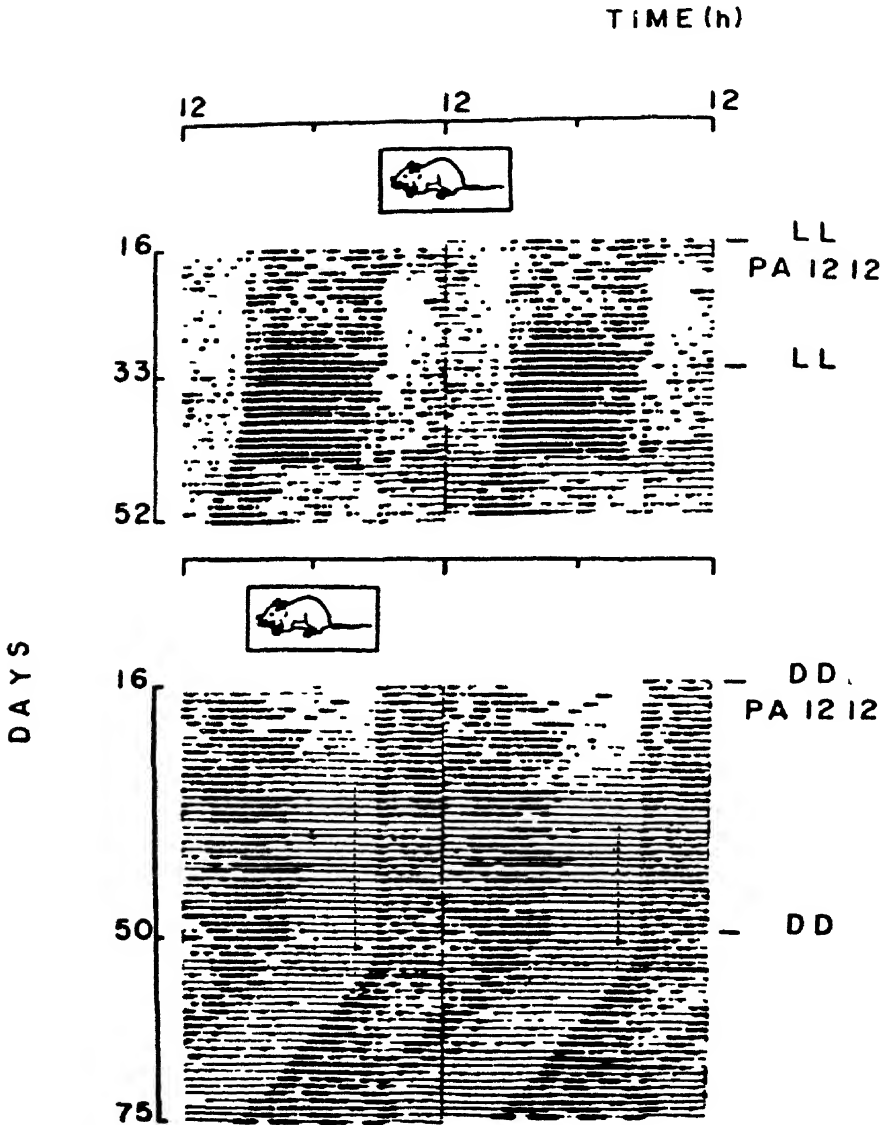


FIG 4 Double-plotted wheel-running activity of the mouse pups. top: recording was done with the PA cycles of mother 12:12 h from day 1 to 18 under the background of continuous illumination(LL) of 0.5 - 0.7 lux. Presence of the mother is from 06.00 h to 18.00 h. The PA cycles were stopped from day 19 onwards and the recording was continued in LL.

because activity is inhibited by the mother's presence and enhanced by her absence (masking). We performed experiments with the period of the presence/absence cycles ranging from 20 to 28 h and find that only periods of 23-25 h allow entrainment and periods below 23h and above 25h do not allow entrainment. Our results speak against the involvement of masking and in favour of the involvement of a genuine circadian organization (Fig. 5).



# HUMAN CIRCADIAN RHYTHMS RECORDED UNDER PROLONGED SOCIAL AND TEMPORAL ISOLATION

The experiments reported here were carried out in a specially built isolation facility for the study of human circadian rhythms which was funded by the award of a "Unit of Neurobiology and Mechanisms of

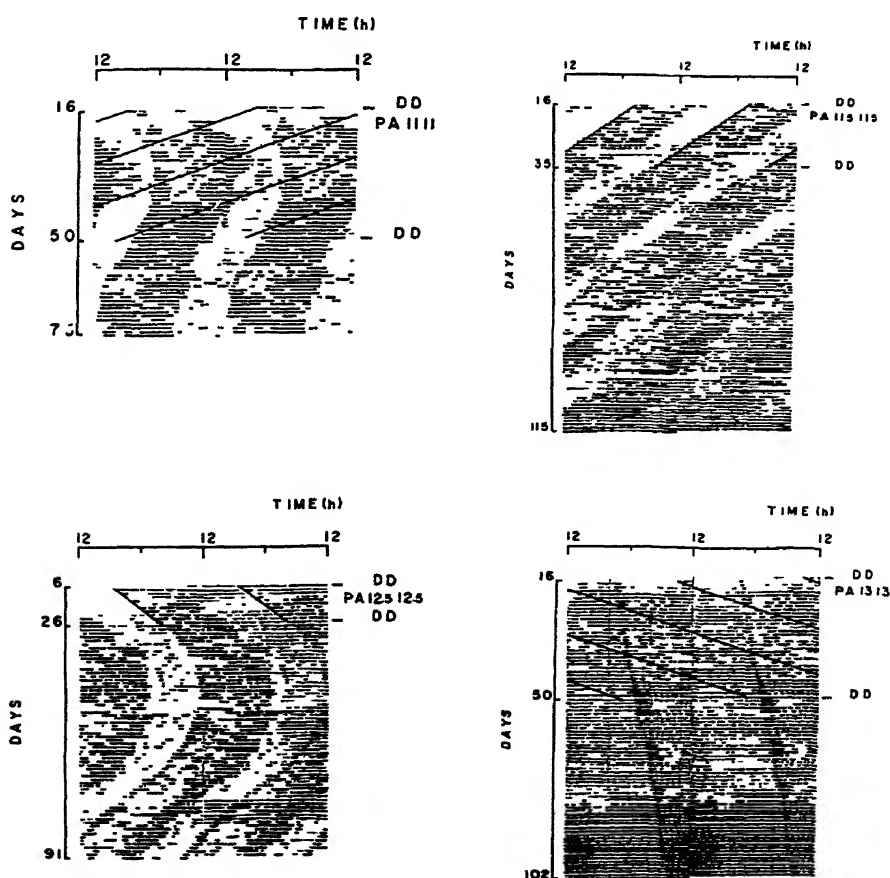


FIG 5 Double-plotted wheel-running activity of mouse pups recorded at different PA cycles in DD. T = 22 h (to left), T = 23 h (top right), T = 25 h (bottom left) and T = 26 h (bottom right). The lines drawn across the recordings indicate the beginning of the absence of the mother. The pups entrain to PA cycles of 11.5:11.5 h and 12.5:12.5 h only and not to 11:11 h and 13:13 h. (After ref. 13).

Behaviour" 1983-88 by the Department of Science and Technology under their IRHPA scheme. There are five other such facilities in Germany, U.K., U.S.A., Japan and Switzerland.

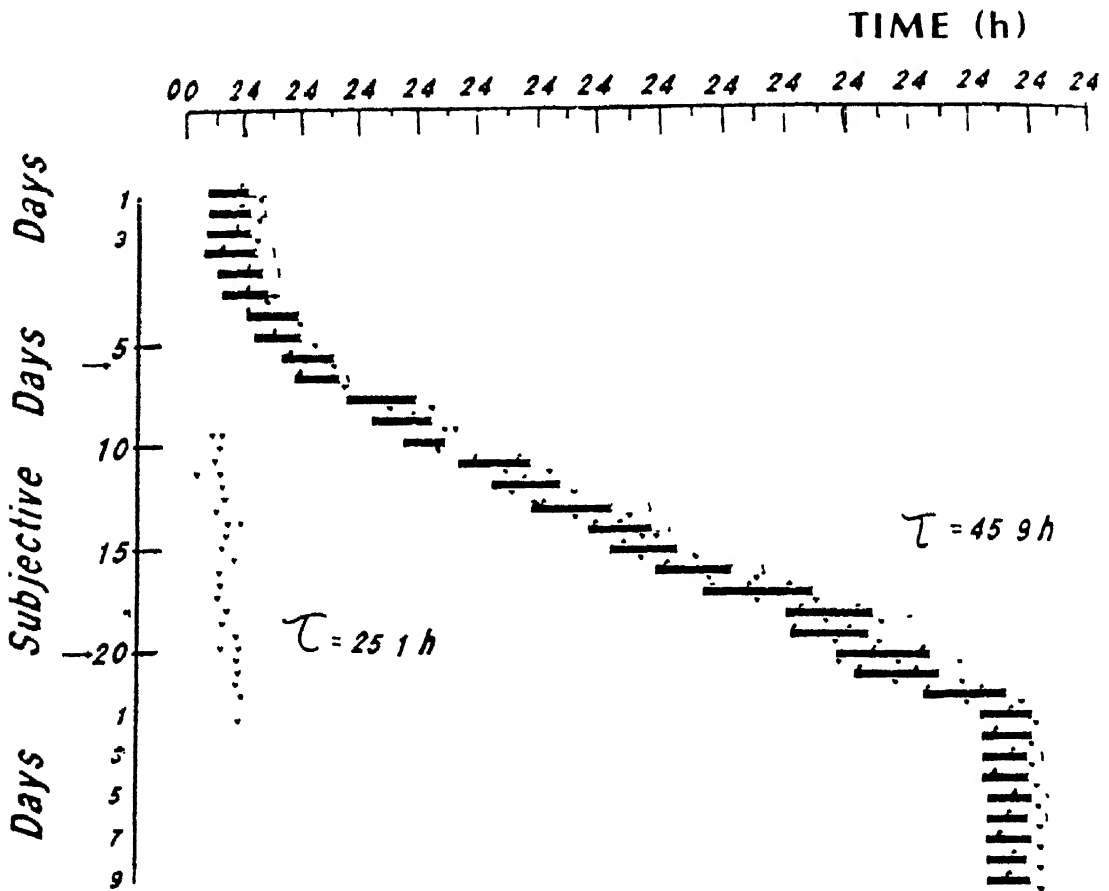


FIG 6 Circadian and circabidian rhythms of sleep and wakefulness (white and black bars) and of rectal temperature (maxima and minima) of a 24-year-old female human subject living in social and temporal isolation for a period of 35 calendar days. Data for days 1-3 before isolation and for days 1-9 after isolation are depicted as calendar (24 h) days. The experiment lasted a total of (3+35+9=) 47 days. During her social and temporal isolation of 35 calendar days the subject experienced only 22 subjective SW days. The Temp rhythm desynchronized from the SW rhythm on subjective day 11. The upper and lower arrows indicate onset of the first and second episodes of menses on subjective days 6 and 20 respectively. Data on the SW rhythm during isolation are plotted in 'subjective days' whereas the Temp rhythm is plotted on both calendar- and subjective-day scales. (After ref. 14).

The isolation facility consisted of a double walled bunker impervious to natural light and external noise. It had a large living area, a kitchenette and a bath. Fluorescent tubelights constituted the light source. The temperature was held constant around 25°C and stored water of uniform temperature was available for use. The isolation facility was devoid of potential zeitgebers, viz, clocks, radios, TV, current periodicals, etc. The occasional food and other requirements of the subject were placed in an antechamber, and there was no social contact for the duration of the experiment. Communication with the outside was mostly through scribbled notes.

We investigated the circadian rhythm in sleep-wakefulness (SW rhythm) and in the rectal-temperature profile (Temp rhythm) of a 24-year old female subject under conditions of social and temporal isolation. The subject stayed in the isolation facility, for 35 days. In isolation the Temp rhythm of the subject freeran with a period of  $25.1 \pm 0.8$  h, but her SW rhythm freeran with a period of  $45.9 \pm 2.1$  h (circabidian), resulting in desynchronization of the two rhythms. In 35 calendar days the subject experienced only 22 subjective, sleep-wakefulness days. Interestingly the menstrual cycle of this subject was normal, ie. two episodes of onset of menses occurred 28 calendar days apart. We conclude in this first report on the subject that the menstrual cycle in the human female may not be coupled to the circadian rhythm underlying sleep-wakefulness while under social and temporal isolation (Fig. 6).

In spite of three decades of intense experimentation it is still not known how the passage of time is perceived by humans living in social isolation. We have investigated how human subjects in our isolation facility estimated 2 h intervals without the aid of watches or clocks. Our studies reported here impressively confirm the correlation between TE (time estimates) of 2 h and hours of wakefulness ( $\alpha$ ) (Fig 7). We conclude therefore that TE are coupled with the circadian rhythm underlying sleep-wakefulness.

Aschoff (1985) found that short time estimates of 10, 20, 30 and 120 s are independent of the duration of  $\alpha$  and claimed that "... somewhere between 120 s and 1 h the perception of time seems to change from being independent to being dependent on  $\alpha$ ". In contrast, we find a correlation of TE of 1 and 2 min with  $\alpha$  values longer than 20 h (unpublished). It would appear from our findings on TE that it is difficult to precisely measure the interval at which TE switches from independence to dependence on  $\alpha$ . The

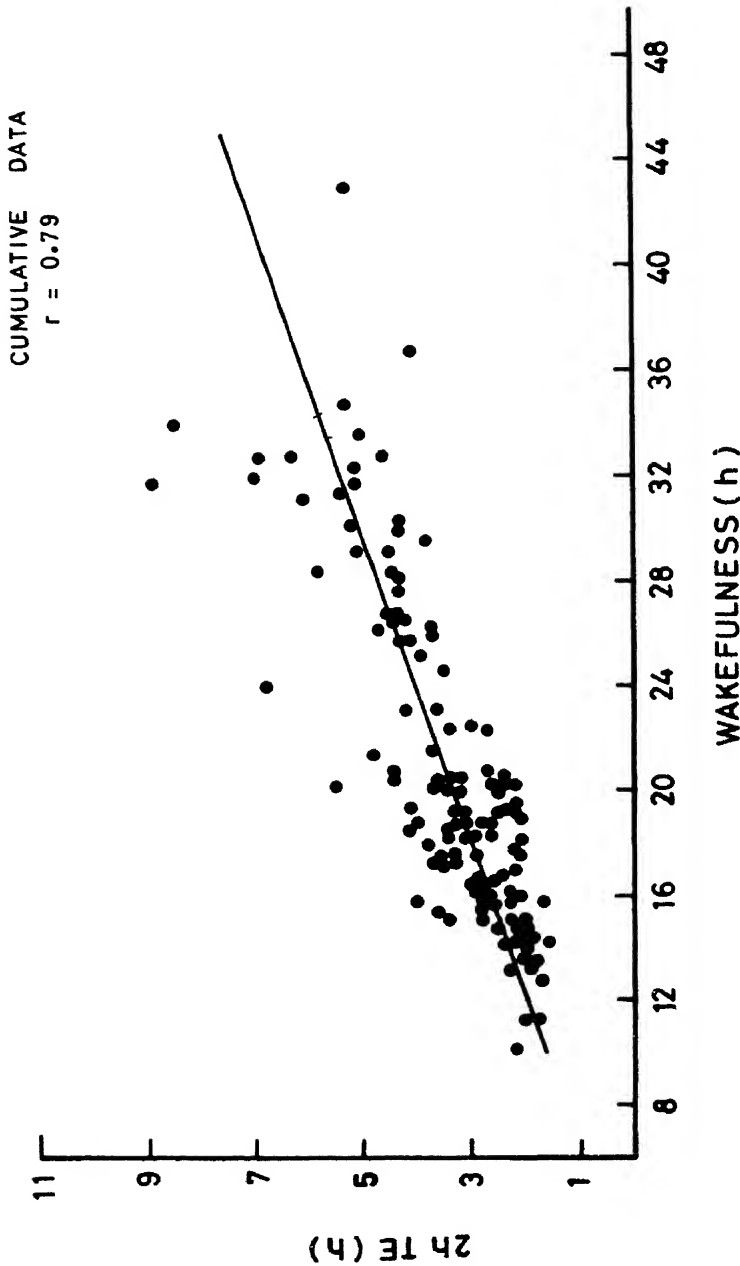


FIG 7 The direct correlation between 2 h TE and  $\alpha$  representing 110 data points made on seven human subjects living in social isolation for periods of 15-35 days. Each point represents the mean estimates of 5-7 readings made during one 'day'.

\*Significant at  $P < 0.01$ . (After ref. 15)

values of  $\alpha$  seem to be more critical in the equation between TE and  $\alpha$ . TE of 2 h of all our subjects show a strong correlation with  $\alpha$ . We are unable to explain our finding that TE is a function of  $\alpha$  in humans in isolation, a finding which suggests that the human mind 'knows' already at the time of waking how long the day will last. The coupling of TE with the  $\alpha$  of the circadian rhythm we are reporting, is reminiscent of the coupling of

ultradian courtship song cycles of ca 1 min in *Drosophila melanogaster* with  $\alpha$  of the circadian rhythm in their locomotory activity.

All subjects adjusted meal timings in consonance with  $\alpha$  so that the midday lunch was roughly midway through  $\alpha$ . Although TE could be made only during  $\alpha$ , the correlation between the two suggests that  $\alpha$  and TE are coupled with the basic circadian rhythm underlying sleep wakefulness. The circadian sleep wakefulness rhythm and the circadian rhythm modulating core temperature dissociated only in one case in our studies. Our own data from desynchronized rhythms clearly demonstrate that TE are not related to the circadian clock that controls the rhythms of automatic functions such as body temperature. We reported in a recent communication<sup>(14)</sup> the menstrual cycle in a human female in isolation, contrary to TE was not correlated with the circadian rhythm in sleep-wakefulness (Fig. 6). These findings lend credence to models which postulate that the sleep-wakefulness rhythm and the body temperature rhythm are driven by two separate clocks. We have preliminary data which leads us to surmise that the coupling of TE with  $\alpha$  may be obtained only under conditions of physiological and physical harmony, and that under stressful conditions leading to discomfort and fever the coupling between the two might break down.

### ACKNOWLEDGEMENTS

I am grateful to a series of the brightest Ph.D. scholars I have had the good fortune to groom since 1975 and in the process learn many new facts of chronobiology myself. Of these joint adventures I would like to make particular mention of Dr. R. Subbaraj, Dr. G Marimuthu, Dr. Dilip Joshi, Dr. N. Viswanathan and Miss L. Geetha for much help, co-operation and unbounded enthusiasm. I am grateful to the D.S.T. for supporting my researches over the years.

I am grateful to the Indian National Science Academy for the award of the Jawaharlal Nehru Birth Centenary Lectureship 1991 which is an honour of the highest order. I also thank Dr. D. Balasubramanian FNA for organizing the lecture in the CCMB.

### References

1. Buening E Opening address: Biological clocks. *Cold Spring Harb Symp quant Biol* 25 (1960) 1-9.

2. Buening E Circadian rhythms, light and photoperiodism : A re-evaluation *Bot Mag Tokyo* 92: (1979) 89-103
3. Subbaraj R and Chandrashekar MK Rigid internal timing in the circadian rhythm of flight activity in a tropical bat *Oecologia (Berl)* 29 (1977) 341-348
4. Halberg F, Visscher MB and Bitner JJ Relations of visual factors to eosinophil rhythm in mice *Am J Physiol* 179 (1954) 229-235
5. Chandrashekar MK Social cues and circadian rhythms. *Curr Sci* 51 (1982) 158-167
6. Marimuthu G, Rajan S and Chandrashekar MK Social entrainment of the circadian rhythm in the flight activity of the microchiropteran bat *Hipposideros speoris Behav Ecol Sociobiol* 8 (1981) 147-150
7. Marimuthu G, Subbaraj R and Chandrashekar MK Social synchronization of the activity rhythm in cave dwelling insectivorous bat *Naturwissenschaften* 65 (1978) 600
8. Marimuthu G and Chandrashekar MK Continuous light inside a cave abolishes the social synchronization of the circadian rhythm in a bat *Behav Ecol Sociobiol* 12. (1983) 321-323
9. Marimuthu G and Chandrashekar MK Social cues of a hipposiderid bat inside a cave fail to entrain the circadian rhythm of an emballonurid bat *Naturwissenschaften* 70 (1983) 620
10. Marimuthu G and Chandrashekar MK Communication and synchronization of circadian rhythms in insectivorous bats *Proc Indian Acad Sci (Anim Sci)* 94 (1985) 655-665
11. Viswanathan N and Chandrashekar MK Cycles of presence and absence of mother mouse entrain the circadian clock of pups *Nature (London)* 317 (1985) 530-531
12. Viswanathan N and Chandrashekar MK Continuous light abolishes the maternal entrainment of the circadian rhythm of pups in the field mouse *Experientia* 43 (1987) 576-578
13. Viswanathan N and Chandrashekar MK Limits of maternal entrainment of the activity rhythm in the field mouse *Mus booduga J Comp Physiol A* 163 (1988) 237-242
14. Chandrashekar MK, Geetha L, Marimuthu G, Subbaraj R, Kumarasamy P and Ramkumar MS The menstrual cycle in the human female is not coupled to the circadian rhythm of sleep-wakefulness in social isolation. *Curr Sci* 60 (1991) 703-705
15. Chandrashekar MK, Marimuthu G, Subbaraj R, Kumarasamy P, Ramkumar MS and Sripathi K Direct correlation between the circadian sleep-wakefulness rhythm and time estimation in humans under social isolation. *J Biosci* 16 (1991) 97-101



**Narasimhaiengar Mukunda** (b. 25 January 1939) did Ph.D. (1964) from University of Rochester, USA. He is Professor, Centre for Theoretical Studies, Indian Institute of Science, Bangalore (1972- ); and was Jawaharlal Nehru Fellow.

Mukunda has contributed to representation theory of Lorentz, Poincare and other non-compact groups; majorana and other relativistic wave equations, dynamics and symmetry of constrained systems; coherent states and correlation functions in quantum optics; Fourier optics and theory of optical systems for the Maxwell field; classical dynamical

formalism; Hamiltonian theories of interacting relativistic particles, use of fibre bundles and differential geometric methods in mechanics and in modern monopole theory; and theories of relativistic particles with internal structure

Mukunda is Fellow, Indian Academy of Sciences, Member, Indian Physics Association and Indian Association for General Relativity and Gravitation, has been Member of INSA Council (1987-89) He is the recipient of Gold Medal (University of Delhi) (1958); Shanti Swarup Bhatnagar Prize (1980), Meghnad Saha Award (Hari Om Ashram Trust) (1985). Jawaharlal Nehru Birth Centenary Lecture Award (INSA) (1992).

*Narasimhaiengar Mukunda was elected to the fellowship of the Academy in 1980*

**Address:** Professor, Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560 012; (080) 344411/2266; Telex 0845-8349 IISc IN; Fax. 080-341683; Res: 103, 6th Main, Malleswaram Bangalore 560 003, Tel. (080) 348657

# PHASE SPACE METHODS IN QUANTUM MECHANICS AND OPTICS

N MUKUNDA FNA

*An account is given of the many ways in which methods based on phase space and the real symplectic groups can be fruitfully exploited in quantum mechanics and optics. These include the concept of generalised rays and action of optical systems on them; the properties of variance matrices for general  $n$ -dimensional quantum systems; a coset-space analysis of Gaussian pure states; a global approach to the concept of squeezing for multimode systems, and an intriguing instance of complementarity between Bose statistics for generalized rays and action of optical systems on them.*

## INTRODUCTION

I would like to describe here some work done over the past few years in the general areas of quantum mechanics and optics, exploiting phase space methods and properties of the family of real symplectic groups. As it will become clear as we go along, it has been a case of successful transfer of ideas and techniques from one area to another. Anyone working in theoretical high energy physics or related fields gets acquainted with a variety of methods based on mechanics, group theory, operator techniques and ideas of symmetry. We have found it possible to exploit such methods in theoretical optics quite effectively, and it is always a pleasure to see that what one has learnt in one field can be used to illuminate another.

Let me begin with a brief historical introduction. As is well known, when one makes the transition from classical to quantum mechanics, classical commuting dynamical variables get replaced by noncommuting quantum mechanical operators. One can ask whether it is possible to set up in any reasonable way a one-to-one correspondence between these two sets of quantities. It turns out that there is no natural, intrinsic or preferred correspondence of this kind, but there are several, each based on some convention. The most convenient of these, in many ways, is that due to Weyl, which was suggested by him already around 1930<sup>1</sup>. Let us recall the details by looking first at systems with one degree of freedom. For such a classical system we have a two-dimensional phase space with basic variables  $q$  and  $p$ , and a general classical dynamical variable is some



numerical function of them. For a quantum system with one degree of freedom, we have operators  $\hat{q}$  and  $\hat{p}$  obeying the Heisenberg commutation relation

$$[\hat{q}, \hat{p}] = i\hbar \quad (1)$$

and a general dynamical variable is an operator function of these basic ones. The Weyl rule of association is completely specified by saying that it is linear and that, for all real numbers  $\lambda$  and  $\mu$  we have the correspondence:

$$\begin{aligned} \text{classical variable } e^{i(\lambda q - \mu p)} &\rightarrow \\ \text{quantum operator } e^{i(\lambda \hat{q} - \mu \hat{p})} &\dots(2) \end{aligned}$$

Therefore, for a general classical quantity  $f(q, p)$  the quantum operator image is given by performing a Fourier transformation:

$$\begin{aligned} f(q, p) &= \int d\lambda d\mu \tilde{f}(\lambda, \mu) e^{i(\lambda q - \mu p)} \rightarrow \\ \hat{F} &= \int d\lambda d\mu \tilde{f}(\lambda, \mu) e^{i(\lambda \hat{q} - \mu \hat{p})} \dots(3) \end{aligned}$$

A nice feature of this Weyl rule is that real classical quantities get mapped to hermitian quantum operators. Of course, from  $\hat{F}$  one can recover  $f(q, p)$

Somewhat later, in 1932, Wigner introduced an interesting way to represent states of a quantum system, somewhat different from their description using wave functions<sup>2</sup>. This was in the context of studying departures of quantum statistical mechanics from classical statistical mechanics. While the representation of quantum states *via* state vectors or wave functions emphasizes the superposition principle of quantum mechanics, the Wigner representation is at the level of the density operator, and so the superposition principle is in a sense hidden and not manifest. Again for a system of one-degree of freedom if the wave function  $\psi(q)$  is known, then the Wigner distribution is a function  $W(q, p)$  on the classical phase space defined by

$$W(q, p) = \frac{1}{2\pi\hbar} \int dq' \psi\left(q - \frac{1}{2}q'\right) \psi^*\left(q + \frac{1}{2}q'\right) e^{ipq'/\hbar} \dots(4)$$

It is easy to see that this is a real function on phase space, but it is not always or everywhere non-negative in general. And what ties together the Weyl and Wigner associations is this: if a state  $|\psi\rangle$  and a dynamical variable

$\hat{F}$  are given for a quantum mechanical system, and the classical variable corresponding to  $\hat{F}$  is the function  $f(q,p)$ , then

$$\langle \psi | \hat{F} | \psi \rangle = \int dq dp f(q,p) W(q,p) \quad \dots(5)$$

So quantum expectation values appear like classical phase space averages.

One can appreciate that even the noncommutative multiplication of quantum operators, and the taking of their commutators, can be given a classical phase space description. This was actually done in explicit form by Moyal some years later<sup>3</sup>, so we shall refer to this entire technology or machinery as the WWM method in quantum mechanics.

One important feature of all this which, however, seems not to have been so clearly expressed, or at least not much emphasized, is the very special role played by the family of real symplectic groups  $Sp(2n, R)$ <sup>4</sup>. This group already appears when one has one degree of freedom,  $n=1$ . In most applications of group theory to problems of symmetry in quantum mechanics, the groups one often encounters are the unitary or real orthogonal ones,  $SU(n)$  or  $SO(n)$ . As a result most of us have an intuitive appreciation or understanding of their structures and properties. In comparison the symplectic groups seem to have some strangeness about them. This is rather unfortunate since they are so closely connected to the basic structures of phase space and canonical Hamiltonian mechanics, which are at the base of both classical and quantum mechanics. I shall introduce and briefly describe this family of groups, and the connection to the WWM method, after generalising the latter to an arbitrary number,  $n$ , of degrees of freedom.

To simplify matters hereafter let us set  $\hbar = 1$  when dealing with a  $2n$ -dimensional phase space and canonical coordinates  $q_r, p_r$ ,  $r=1, 2, \dots, n$ . To deal with them and express Poisson Brackets (PB) in a compact manner we collect them into a  $2n$ -component column vector ( $\xi_a$ ):

$$\xi = (\xi_a) = \begin{pmatrix} \cdot \\ \cdot \\ \xi_1 \\ \cdot \\ \xi_2 \\ \vdots \\ \xi_{2n} \end{pmatrix} = \begin{pmatrix} q_1 \\ \dot{q}_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix} \quad \dots(6)$$

We list the  $q$ 's first and then the  $p$ 's. The basic PB's can be written as

$$\{\xi_a, \xi_b\} = \beta_{ab}, \beta = \begin{pmatrix} 0 & 1n \times n \\ -1n \times n & 0 \end{pmatrix} \quad (7)$$

This matrix  $\beta$ -antisymmetric and even dimensional — will turn out to be fundamental for  $Sp(2n, R)$  as a "symplectic metric". For the quantum mechanical case we have operators  $\hat{q}_r, \hat{p}_r$  put together to form a column vector  $\xi$  obeying canonical commutation relations:

$$\hat{\xi} = \begin{pmatrix} \hat{q}_1 \\ \hat{q}_n \\ \hat{p}_1 \\ \vdots \\ \hat{p}_n \end{pmatrix} \quad [\hat{\xi}_a, \hat{\xi}_b] = i\beta_{ab} \quad (8)$$

Now let  $\hat{\Gamma}$  be any operator pertaining to the quantum system. It could for instance be the density operator for a pure or mixed state. We define the Wigner distribution corresponding to  $\hat{\Gamma}$  to be a function on the corresponding classical phase space by

$$W(\xi) = (2\pi)^{-n} \int d^n q' \left\langle \underline{q} - \frac{1}{2} \underline{q}' \right| \hat{\Gamma} \left| \underline{q} + \frac{1}{2} \underline{q}' \right\rangle e^{i \underline{p} \cdot \underline{q}'} \quad (9)$$

It is simple to check that one can recover the coordinate-space kernel of  $\hat{\Gamma}$  from  $W(\xi)$ ; so  $\hat{\Gamma}$  and  $W(\xi)$  contain the same information. Then the  $n$ -dimensional version of (5) holds and we may express it as

$$Tr(\hat{\Gamma} \exp\{i(\underline{\lambda} \cdot \hat{\underline{q}} - \underline{\mu} \cdot \hat{\underline{p}})\}) = \int d^{2n} \xi W(\xi) \exp\{i(\underline{\lambda} \cdot \underline{q} - \underline{\mu} \cdot \underline{p})\} \quad \dots(10)$$

By linearity and Fourier transformation this kind of relation holds for expectation values of general quantum mechanical operators.

The  $2n$ -dimensional real symplectic group  $Sp(2n, R)$  can now be introduced. It consists of all real linear invertible transformation on the  $q$ 's and  $p$ 's such that classically the PB's and quantum mechanically the commutation relations are preserved:

$$\xi'_a = S_{ab} \xi_b : \{\xi'_a, \xi'_b\} = \beta_{ab},$$

and

$$\hat{\xi}'_a = S_{ab} \hat{\xi}_b : [\hat{\xi}'_a, \hat{\xi}'_b] = i\beta_{ab}, \quad \dots(11)$$

The condition on  $S$ -the defining property of  $Sp(2n, R)$ -is

$$S \in Sp(2n, R) \Leftrightarrow S\beta S^T = \beta \quad \dots(12)$$

In other words the symplectic metric must be preserved. Because  $\beta$  is antisymmetric non-singular, these groups exist in even dimensions only.

The preservation of the  $PB$ 's or commutation relations, as the case may be, means that classically we have a canonical transformation, and in quantum mechanics a unitary one. The latter means that for any symplectic matrix,  $S$ , we have a unitary operator  $U(S)$  which implements the above transformation through conjugation:

$$S \in Sp(2n, R): U(S)^{-1} \hat{\xi} U(S) = S \hat{\xi} \\ U(S_1) U(S_2) = U(S_1 S_2) \quad \dots(13)$$

There is an unavoidable sign ambiguity here, similar to the sign ambiguity in the spin half representation of angular momentum. So one says that one has here a faithful representation, not of  $Sp(2n, R)$  but of its two-fold covering metaplectic group  $Mp(2n)^5$ .

Now comes the fundamental connection to and significance of the Wigner distribution, indeed of the entire WWM technique. It is that when an operator  $\Gamma$  undergoes the operator symplectic transformation, its Wigner function changes in a very simple way by a point transformation in phase space:

$$\hat{\Gamma} = U(S) \Gamma U(S)^{-1} \leftrightarrow W'(\xi) = W(S^{-1}\xi) \quad \dots(14)$$

Indeed one can view the WWM procedure as a way to make this operator action of  $Sp(2n, R)$  as simple and classical looking as possible. We can take this as the defining characteristic of this procedure. The infinitesimal aspect of this relationship is also worth exhibiting. Since we are considering linear canonical transformations, the generators are quadratic in the  $q$ 's and  $p$ 's. That is, the unitary operator  $U(S)$  is essentially the exponential of a symmetric quadratic expression in the operators  $\hat{q}, \hat{p}$ . Pursuing this, one finds that the processes of taking the anticommutators or commutators of an

operator  $\hat{\Gamma}$  with the  $\hat{q}$ 's and  $\hat{p}$ 's has a nice appearance in terms of the Wigner function<sup>6</sup>.

$$\frac{1}{2} \{ \hat{\xi}_a, \hat{\Gamma} \} \leftrightarrow \xi_a W(\xi),$$

and

$$\{ \hat{\xi}_a, \hat{\Gamma} \} \leftrightarrow i\beta_{ab} \frac{\partial}{\partial \xi_b} W(\xi), \quad \dots(15)$$

It is this intimate relationship between the symplectic groups and the WWM method in quantum mechanics that somehow seems to have been not as much emphasized and exploited as one might have expected. Already for one degree of freedom,  $n = 1$ , we have something nontrivial. The group  $Sp(2, R)$  is easy to define as a group of real  $2 \times 2$  unimodular matrices:

$$Sp(2, R) = \left\{ S = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid ad - bc = 1 \right\} \quad \dots(16)$$

And there are many uses of it in various contexts.

After this brief exposure to these concepts, let me turn for a while to some problems in statistical optics. As is well known, for the theoretical treatment of all classical optical experiments, and for a first understanding of coherence and partial coherence, the basic object is the so-called two-point correlation function whose study in a systematic way was pioneered by Wolf long ago<sup>7</sup>:

$$\Gamma(x; x') \sim \langle A(x) A(x')^* \rangle \quad \dots(17)$$

I leave unspecified the dimensionality of the spatial or space-time arguments  $x$  and  $x'$ . Here,  $A(x)$  is the vector potential or the electric field as one wishes; more precisely it is the positive frequency (analytic signal) part of the full real field. And the angular brackets denote an ensemble or statistical average. This two-point function can be defined either classically or quantum mechanically, though at this level one cannot distinguish between the two theories. The intensity of light at a spacetime point is a particular case of this function, when the arguments coincide. And what we mean by

classical optical experiments are all those which only involve measurements of intensity of light at one space-time point at a time. The two-point function is hermitian in the obvious sense, and regarded as the kernel of an operator it is also positive semi-definite.

Now Wolf in 1976 posed the following problem<sup>8</sup>: is there a way to obtain the phenomenological quantities and equations of the theory of radiative transfer from a starting point based on Maxwell's field equations and statistical ideas? He came to the conclusion that this was possible provided one worked with the 2-point function, and the ensemble was time-stationary. The latter means that one can deal with radiation of different frequencies independently. However, Wolf's original prescription for going from the two-point function to the quantities that arise in radiative transfer theory turned out to be not very convenient, and a much more elegant prescription was soon given by Sudarshan in 1979<sup>9</sup>. This was just a transcription of the entire WWM technology from quantum mechanics to optics! I will describe it now in the context of paraxial optics<sup>10</sup> — this is the situation where one has a beam of light propagating along a given spatial direction, say the  $z$ -axis; and one is interested in the properties across any transverse plane, and how they develop during propagation.

Let then  $\Gamma(x_{\perp}, x'_{\perp})$  denote the two-point correlation function of a beam across a transverse plane, so  $x_{\perp}$  and  $x'_{\perp}$  are two-dimensional vectors, and the common coordinate  $z$  is suppressed. So except for the normalization condition it is just like the density matrix for a two-dimensional quantum system. In optics, one has the total beam intensity given by

$$I = \int d^2 x_{\perp} \Gamma(x_{\perp}; x_{\perp}) \quad \dots(18)$$

whereas in quantum mechanics one would have demanded that this integral be unity. Now the WWM transform of  $\Gamma(x_{\perp}; x'_{\perp})$  is called the Wolf function or generalized ray distribution function  $W(x_{\perp}; p_{\perp})$ :

$$W(\xi) = W(x_{\perp}; p_{\perp}) = (2\pi)^{-2} \int d^2 x'_{\perp} \Gamma\left(x_{\perp} - \frac{1}{2}x'_{\perp}; x_{\perp} + \frac{1}{2}x'_{\perp}\right) e^{ip_{\perp} \cdot x'_{\perp}},$$

$$\xi = \begin{pmatrix} x_{\perp} \\ p_{\perp} \end{pmatrix} \quad \dots(19)$$

One says that  $W(x_{\perp}; p_{\perp})$  is the "intensity" of generalized rays of light at transverse position  $x_{\perp}$  and transverse direction  $p_{\perp}$ . We must remember of course that while  $W(\xi)$  is definitely real, it could sometimes be negative. So in the terminology introduced by Sudarshan<sup>9</sup>, when  $W(\xi)$  is positive we say we have bright rays, and when it is negative we say we have dark rays. But the true justification for the use of the name "generalised rays" really comes from the connection to the symplectic groups. Let me explain.

In ray optics proper there is a class of optical systems called First Order Optical Systems—FOS for short—defined by the way they act on the position and direction coordinates of geometrical rays<sup>11</sup>. Remembering the paraxial limit, an incoming ray is specified by the four variables  $\xi = (x_{\perp}, p_{\perp})$  of transverse position and direction. FOS's stand in one-to-one correspondence with matrices  $S \in Sp(4, R)$  and act on rays in this way.

*FOS Action on Geometrical Rays*

$$\begin{pmatrix} x_{\perp} \\ p_{\perp} \end{pmatrix}_{in} \rightarrow \begin{pmatrix} x'_{\perp} \\ p'_{\perp} \end{pmatrix}_{out} = S \begin{pmatrix} x_{\perp} \\ p_{\perp} \end{pmatrix}_{in}, S \in Sp(4, R) \quad . \quad (20)$$

One may ask why the symplectic nature comes in here at all. The reason is beautiful and deep: it is because ray optics is governed by Fermat's principle, and this is a variational or extremum principle just as in Lagrangian mechanics. So the entire canonical structure and machinery is automatically there and one cannot escape it! Now when one goes to wave optics (let us ignore polarization here), there is a wave amplitude  $\psi(x_{\perp})$  over any transverse plane, and the passage of the wave through any FOS is given by the action of an integral kernel on the amplitude: it is just the action by the metaplectic unitary operator  $U(S)$  encountered earlier!

*FOS Action on Waves*

$$\begin{aligned} \psi' &= U(S) \psi : \\ \psi'(x'_{\perp}) &= \int d^2 x_{\perp} \langle x'_{\perp} | U(S) | x_{\perp} \rangle \psi(x_{\perp}) \end{aligned} \quad \dots(21)$$

I am using here the suggestive and natural notation of quantum mechanics in two dimensions. This integral kernel  $\langle x'_{\perp} | U(S) | x_{\perp} \rangle$  is called the generalized Huyghens kernel, and is the exponential of a quadratic

expression in  $x'_\perp$  and  $x_\perp$ <sup>12</sup>. What the definition of the generalized ray distribution function achieves is that this action of *FOS's* on waves gets reduced to the kind of action we had on geometrical rays, but now in the context of a full wave theory. We go from  $\psi$  to  $\Gamma$  and then on to  $W$  to find:

*FOS Action on Generalized Rays*

$$W'(\xi) = W(S^{-1}\xi) \quad \dots(22)$$

With our improved appreciation of the significance of the symplectic groups, we can in fact regard this simple action of *FOS's* as the prime motivation for introducing the Wolf function and creating the concept of generalized rays rather than as a bridge to the concepts of radiative transfer theory.

For axially symmetric optical systems we have a simplification from the group  $Sp(4, R)$  to  $Sp(2, R)$ . I might indicate some examples of physical systems, and give their corresponding matrix representatives in this case<sup>11</sup>:

$$\text{Free propagation over distance } D: f(D) = \begin{pmatrix} 1 & D \\ 0 & 1 \end{pmatrix}$$

$$\text{Lens of power } g: l(g) = \begin{pmatrix} 1 & 0 \\ -g & 1 \end{pmatrix}$$

$$\text{Magnification: } m(\xi) = \begin{pmatrix} e^\xi & 0 \\ 0 & e^{-\xi} \end{pmatrix}$$

$$\text{Fourier transformation: } \mathcal{F} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \dots(23)$$

The key point is that the action of integral transformations on the wave amplitude gets reduced to simple point transformations on generalized rays. The rays are *generalized*, not geometrical, because we are presenting here a complete wave theory, not a short wave length approximation to it!

Now I would like to outline some further uses of the symplectic groups in quantum mechanics and optics, in conjunction with the WWM method. We shall be concerned with general  $n$ -dimensional quantum



systems-paraxial optics corresponding to  $n=2$ . For any quantum state (or optical beam) the variances or spreads in the basic coordinates and momenta contain a certain amount of important information about the state of the system. If the state, i.e., the two-point function or density matrix or Wigner-Wolf function, is furthermore Gaussian, these variances essentially determine the state completely (the additional information being the means). In general, for any state, we define the real symmetric  $2n$ -dimensional variance matrix  $V$  as

$$V = (V_{ab}),$$

$$V_{ab} = \frac{1}{2} \text{Tr} \hat{\Gamma} \{ \hat{\xi}_a - \xi_a^{(0)}, \hat{\xi}_b - \xi_b^{(0)} \} / \text{Tr} \hat{\Gamma} \quad \dots(24)$$

where  $\xi_a^{(0)}$  are the mean values of the  $\hat{\xi}_a$ . In terms of the Wigner-Wolf function this has a neat form

$$V_{ab} = \int d^{2n} \xi \xi_a \xi_b W(\xi) / \int d^{2n} \xi W(\xi) - \xi_a^{(0)} \xi_b^{(0)} \quad \dots(25)$$

Since the means are easy to handle, I omit them hereafter. Under action by any symplectic transformation (FOS in optics) this matrix has a nice law of change:

$$\hat{\Gamma}' = U(S) \hat{\Gamma} U(S)^{-1} \rightarrow V' = S V S^T \quad \dots(26)$$

If we assume that we have a Gaussian phase space distribution, the knowledge of  $V$  fixes  $W(\xi)$  completely:

$$\text{Gaussian distribution : } W(\xi) \simeq \exp \left( -\frac{1}{2} \xi^T V^{-1} \xi \right) \quad \dots(27)$$

Certain properties of the variance matrix are evident upon inspection — these are that it has to be real symmetric positive definite. But not every such matrix qualifies to be the variance matrix of some physical state! There are certain inequalities that  $V$  must obey in order to be physically allowed. It is these that ensure that when we go back from  $W(\xi)$  to the coordinate space kernel of  $\hat{\Gamma}$ , then this operator is positive semidefinite. In quantum mechanics these are the usual Heisenberg uncertainty relations, or rather extensions of them taking into account all the elements of the variance matrix. In optics there are similar uncertainty relations for the positions and directions of hypothetical rays in wave optics.

One can ask if there is a way to state these conditions on  $V$  in an explicitly symplectic invariant form. It is interesting that this can be done<sup>13</sup>; the key is a very nice theorem due to Williamson which I would like to describe at this point<sup>14</sup>.

As is well known, any real symmetric matrix can be brought to diagonal form by a suitable real orthogonal transformation, and then the diagonal elements are the eigenvalues of the original matrix. However, under  $Sp(2n, R)$ , the variance matrix transforms by a symmetric symplectic transformation, as in (26), and this is not a similarity transformation. Of course the reality, symmetry and positive definiteness of  $V$  are maintained. Now the theorem of Williamson says: if  $V$  is symmetric positive definite, we can definitely find a symplectic matrix  $S$  such that  $SVS^T$  is diagonal. Of course the resulting diagonal elements will in general not be eigenvalues of  $V$ . (Notice by the way that if  $V$  is not symmetric positive (or negative) definite, it may not be possible to "diagonalise" it with a symplectic  $S$ ). the complete set of inequalities on  $V$  expressing the uncertainty principles in invariant form are just these: with  $S$  chosen such that  $SVS^T$  is diagonal, the elements of  $V'$  must obey the ordinary Heisenberg relations for each degree of freedom:

$$\Delta q'_r \Delta p'_r \geq \frac{1}{2}, r = 1, 2, \dots, n. \quad \dots(28)$$

These then ensure that  $\hat{I}$  as an operator is nonnegative. It is possible to express these inequalities directly in terms of  $V$ , without having to pass to its Williamson diagonal form at all<sup>13</sup>. It involves working with the traces  $T_r(\beta V)$ <sup>21</sup> for  $l = 1, 2, \dots, n$ . which are individually  $Sp(2n, R)$  invariant. But I will omit the details.

The Gaussian pure states of  $n$ -dimensional quantum systems have a very interesting group theoretical structure which only the combination of phase space and symplectic methods can reveal in a reasonably straightforward way. Let me just indicate how it goes<sup>15</sup>. Such states are described by normalised complex Gaussian wave functions:

$$\psi(q) \simeq \exp \left\{ -\frac{1}{2} q^T (u + iv) q \right\} \quad \dots(29)$$

So apart from a phase each such state is specified by a pair of real symmetric  $n \times n$  matrices  $(u, v)$ , with the former being positive definite. It now turns out that such states are in one-to-one correspondence with points of the coset space  $Sp(2n, R)/U(n)$ , where  $U(n)$  is the maximal compact subgroup of the noncompact group  $Sp(2n, R)$ . When we act on this  $\psi(q)$  by the unitary operator  $U(S)$  for some  $S \in Sp(2n, R)$ , we get another similar Gaussian pure state; and any two such states can be connected in this way. (We are ignoring here the niceties of the metaplectic group!). The change in  $\psi$  is given as a point transformation on the coset space by a very pretty formula; the point  $(u, v)$  jumps to another point  $(u', v')$  specified as follows:

$$S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(2n, R), \quad \wedge = (iu \quad -v)^{-1}.$$

$$\psi' = U(S)\psi \Rightarrow \wedge' = (A\wedge + B)(C\wedge + D)^{-1} \quad . \quad (30)$$

It would take me too far to present any more details but I thought I should at least mention this result. By the way, for one degree of freedom,  $\wedge$  becomes a complex number in the lower half plane, and we have here a Möbius transformation. So what we have above is the  $n$ -dimensional matrix generalisation of the complex half plane and these transformations, which is quite an intricate structure<sup>16</sup>.

Returning to the properties of variance matrices for general states, the combined use of phase space and group theoretical ideas help us understand in a comprehensive way the notion of squeezing. Both theoretically and experimentally this is of considerable interest these days. While much work has been done, it turns out that the methods I have been describing can teach us something new even for one-dimensional systems. They also show how to define squeezing for systems with  $n$  degrees of freedom.

Recall that for one degree of freedom, in dimensionless form, such as for a single mode of the radiation field, the Heisenberg uncertainty principle says

$$\Delta q \Delta p \geq \frac{1}{2} \quad \dots(31)$$

Here there is no reference to the cross variance  $\Delta(qp)$ . Usually, a state is said to be squeezed if one of the two spreads,  $\Delta q$  or  $\Delta p$ , is less than  $\frac{1}{\sqrt{2}}$ . We

have proposed and analyzed in detail a more satisfactory definition of squeezing, involving the complete  $2 \times 2$  variance matrix, and possessing a reasonable degree of invariance<sup>17</sup>. Our definition is:

$$V = \begin{pmatrix} (\Delta q)^2 & \Delta(qp) \\ \Delta(qp) & (\Delta p)^2 \end{pmatrix}, \det V \geq \frac{1}{4}.$$

Squeezed state  $\Leftrightarrow$  lesser eigenvalue of  $V < \frac{1}{\sqrt{2}}$  ... (32)

Remembering the general  $Sp(2, R)$  transformation law of  $V$ , namely as in (26),

$$S \in Sp(2, R) \quad V \rightarrow V' = SVS^T, \quad \dots (33)$$

we see that our definition of squeezing is invariant under the  $U(1)$  or  $SO(2)$  subgroup of  $Sp(2, R)$ , but of course not under the entire group of symplectic transformations. And one can give good physical arguments why this should be so.

Now for states of  $n$ -dimensional systems, in a similar way, one can justify the search for a definition of squeezed states in terms of the variance matrix, such that the definition is  $U(n)$  invariant, but not  $Sp(2n, R)$  invariant. Here, for  $n \geq 2$ , something essentially new arises. We must remember that  $V$  is a  $2n \times 2n$  matrix, while the subgroup of  $Sp(2n, R)$  of interest here is  $U(n)$ . In one dimension, since  $U(1)$  and  $SO(2)$  are the same, any  $2 \times 2$  variance matrix is diagonalisable by some  $U(1)$  element, so we were able to state our squeezing definition in (32) in terms of the eigenvalues of  $V$ . But for  $n \geq 2$  dimensions, a corresponding statement does not hold. Let us now propose our definition of squeezing in  $n$ -dimensions<sup>16</sup>. Look at the diagonal elements of  $V$ : they are the squares of the uncertainties  $(\Delta q_1)^2, (\Delta q_2)^2, \dots, (\Delta q_n)^2, (\Delta p_1)^2, \dots, (\Delta p_n)^2$ . We shall say a given state is squeezed if any diagonal element of  $V$  is less than  $1/2$ , or if this is so for any  $U(n)$  transform of  $V$ . This is an explicitly  $U(n)$ -invariant definition, and it makes good physical sense. However, it is not expressed in terms of the eigenvalues of  $V$ . Now, in general, a  $2n \times 2n$  variance matrix cannot be brought to diagonal form by action of elements in the  $U(n)$  subgroup of  $Sp(2n, R)$ . What comes as a nice surprise is this: the definition of squeezing we have just given is equivalent to the following:

State with variance matrix  $V$  is squeezed  $\Leftrightarrow$  the smallest eigenvalue of  $V$  is less than  $1/2$ .

I hope I have communicated the sense in which we have something nontrivial here. It is worth stressing that it is experience with uses of phase space and symplectic group ideas that allows us to tackle this problem of squeezing in a kind of global way. It turns out that one can also give a broad and physically meaningful classification of states by classifying their variance matrices into distinct families, along these lines.

The final technical item I would like to present concerns the problem of going beyond the two-point correlation function in optics. We saw that generalized rays were introduced in terms of this function. Now there is an infinite hierarchy of correlation functions, all of which are needed for a complete statistical description<sup>18</sup>. The two-point function is essentially the lowest in this hierarchy. A general correlation function  $\Gamma^{(m,n)}$  of order  $(m,n)$  involves the expectation value of  $m$  factors of the analytic signal  $A$  and  $n$  factors of the complex conjugate  $A^*$ . A careful analysis now shows that only for the diagonal functions  $\Gamma^{(m,m)}$  it is possible to give an alternative account in terms of generalized rays<sup>19</sup>. We are unable to do it for the functions with  $m \neq n$ . So the phase space method in this sense goes only part of the way. All the information contained in the functions  $\Gamma^{(m,m)}$  can be reexpressed via a sequence of phase space functions  $W_m$  in an invertible manner:

$$\Gamma^{(m,m)}(\dots x_{\perp} \dots; \dots x'_{\perp} \dots) \leftrightarrow W_m(\xi^{(1)} \dots \xi^{(m)}) \quad \dots(35)$$

This function  $W_m$  can be viewed as the  $m$ -generalized rays joint distribution function, and it has a simple behaviour under the action of symplectic  $FOS$ 's. Incidentally, the Bose nature of light shows up in a nonlocal correlation among generalized rays. One finds that for any pair of arguments  $\xi^{(j)}$  and  $\xi^{(k)}$  the following integral relation holds for  $W_m$ <sup>19</sup>:

$$W_m(\dots \xi^{(j)} \dots \xi^{(k)} \dots) = \int d^4\xi \exp \{ -i\xi^T \beta (\xi^{(j)} - \xi^{(k)}) \} \\ W_m \{ \dots \frac{1}{2} (\xi^{(j)} + \xi^{(k)} - \xi) \dots \frac{1}{2} (\xi^{(j)} + \xi^{(k)} + \xi) \dots \} \quad \dots(36)$$

This is explicitly covariant under  $Sp(4,R)$  transformations. There is a kind of complementarity here. Either one works in configuration space where the Bose property is simple while action by symplectic  $FOS$ 's is nonlocal; or one works with phase space and generalised rays, which makes  $FOS$  action simple but the Bose nature nonlocal. All this goes through with essentially

no changes for the hierarchy of quantum higher order correlation functions as well.

Let me now conclude with a few remarks. I hope to have shown you how the use of phase space methods and symplectic group structure has led to new insights and points of view in optics. These are techniques learnt and perfected in the arena of quantum mechanics, and then successfully transferred to another field. From the epistemological point of view the following rather interesting comparison can be made. In quantum mechanics the meaning of the uncertainty principle, and the problem of describing localisable particles in terms of probability amplitudes as waves, lead to profound issues of observation and existence. In optics there are similar uncertainty principles for ray variables in the wave theory, but clearly the problems of interpretation are nowhere as deep. However a new criterion makes its appearance—that of convenience of description—and in this context the kind of complementarity I pointed out above comes up. This goes to show that ways of thinking acquired in the understanding and interpretation of quantum mechanics come up and become useful in other areas as well. And it even seems that the subjective criterion of convenience of description is relevant in discerning what is real and what exists. I refer here to the natural question: do generalised rays really exist or are they merely a mathematical artefact and a matter of convenience? I will leave these and related questions to the interested reader to ponder over, and only express the hope that she will find some satisfaction in these ways of looking at an old friend with new eyes.

### ACKNOWLEDGEMENT

It is a pleasure to acknowledge with gratitude the lessons I have learnt in all of the above from R Simon, and in much of it from E.C.G. Sudarshan.

### References

- 1 H Weyl *Theory of Groups and Quantum Mechanics* Dover, New York (1931) 274
- 2 E P Wigner *Phys Rev* **40** (1932) 749
- 3 J E Moyal *Proc Camb Phil Soc* **45** (1949) 99
- 4 See however, V Guillemin and S Sternberg *Symplectic Techniques in Physics* Cambridge University Press (1984)
- 5 A Weil *Acta Math* **III** (1964) 143, see also ref. (4) above

- 6 N Mukunda *Pramana* **11** I (1978); R Simon, E C G Sudarshan and N Mukunda *Phys Rev A* **36** (1987) 3868
- 7 M Born and E Wolf *Principles of Optics* 6th Edition, Pergamon Press, Oxford (1987), Chapter X; L Mandel and E Wolf, *Rev Mod Phys* **37** (1965) 231; J R Klauder and E C G Sudarshan *Fundamentals of Quantum Optics* W A Benjamin Inc New York (1968)
- 8 E Wolf *Phys Rev D* **13** (1976) 869
- 9 E C G Sudarshan *Phys Lett* **73A** (1979) 269
- 10 R Simon *Pramana* **20** (1983) 105; R Simon, E C G Sudarshan and N Mukunda *Phys Rev A* **29** (1984) 3273, **A31** (1985) 2419
- 11 See, for example, O N Stavroudis *The Optics of Rays Wavefronts and Caustics* Academic Press New York (1972)
- 12 See, for example, K B Wolf *Integral Transforms in Science and Engineering* Plenum Press New York (1979)
- 13 R Simon, E C G Sudarshan and N Mukunda *Phys Lett A* **124** (1987) 223; *Phys Rev A* **36** (1987) 3868
- 14 J Williamson *Am J Math* **58** (1936) 141; V I Arnold *Mathematical Methods of Classical Mechanics* Springer-Verlag, New York (1978) Appendix 6
- 15 R Simon, E C G Sudarshan and N Mukunda *Phys Rev A* **37** (1988) 3028 and references therein
- 16 R Simon, and N Mukunda (*in preparation*)
- 17 N Mukunda *Curr Sci* **59** (1990) 1135; B Dutta, N Mukunda, R Simon and Ajay Subramaniam "Squeezed States, photon number distributions, and  $U(1)$  invariance" *J opt Soc Am, B* **(10)**, 253 (1993)
- 18 See, for instance, the second and third of refs. (7)
- 19 N Mukunda, E C G Sudarshan and R Simon *Foundations of Physics* **18** (1988) 277



**Bimla Buti** (b. 19 September 1933) did Ph.D. (1962, from University of Chicago, USA. She is Senior Professor (1982- ), (Dean of Faculty 1982-86) at Physical Research Laboratory Ahmedabad; and was formerly Visiting Professor, University of California Los Angeles (1986-87).

Buti has developed mathematical models for interpreting the intriguing and anomalous phenomena observed in interplanetary as well as cometary and a variety of other space and astrophysical plasmas, for example, coherent radiation models for radio pulsars, auroral and

Uranian emissions, solar radio bursts, anomalous heating, and acceleration of charged particles due to turbulent and chaotic plasma processes. She has edited four books on Plasma Physics.

Buti is Fellow of Third World Academy of Sciences (Trieste), National Academy of Sciences (India) and American Physical Society, President, Plasma Science Society of India (1992-93); Member, Sigma Xi, American Physical Society, American Geophysical Union, Plasma Science Society of India, Astronomical Society of India, and Indian Physics Association; has been Member of INSA Council (1991-93). She is the recipient of Vikram Sarabhai Award for Planetary Sciences (1977); Jawaharlal Nehru Birth Centenary Lecture Award (INSA) (1993)

*Bimla Buti was elected to the fellowship of the Academy in 1981*

**Address:** Senior Professor, Physical Research Laboratory, Ahmedabad 380 009; Tel: (0272) 462129; Telex: 121-6397; Fax: 91-272-460502; E-Mail: buti@prl.ernet.in; Res: D/7, PRL Quarters, Navrangpura, Ahmedabad 380 009; Tel: (0272) 429878.



# COHERENCE IN CHAOS

B BUTI

I feel deeply honoured by this award, for which I would like to express my gratitude to the Indian National Science Academy. I feel proud to be associated, through this lecture, with our first Prime Minister Pandit Jawaharlal Nehru who had a big role to play in the scientific development of the country. He was a great source of encouragement to all the scientists.

## INTRODUCTION

The topic of my lecture, namely Chaos, is truly interdisciplinary. It has applications in physics, chemistry, and life sciences. However, I will confine myself to plasma systems. The study of phenomena of chaos, in plasmas, is being pursued essentially to understand the phenomena of plasma turbulence which is very often observed both in laboratory as well as in natural plasmas.

A chaotic system, by definition, is turbulent whereas turbulent system is basically nonlinear. We have used the techniques of nonlinear dynamics to study the chaotic processes in plasmas. As an illustrative example, we will discuss nonlinear, coherent and chaotic Alfvén waves. The reason for choosing Alfvén waves is simply because nonlinear Alfvén waves have been observed<sup>1</sup> in solar wind, planetary bow shocks, interplanetary shocks, environment of comets etc. Not only the large amplitude Alfvén waves but even the Alfvénic turbulence has been observed in the solar wind<sup>2</sup> as well as in the vicinity of comets<sup>3</sup>.

A nonlinear wave, in general, is equivalent to a nonlinear dynamical system which exhibits the phenomena of chaos. In this lecture, we will discuss how and under what conditions, the nonlinear Alfvén waves can become chaotic; these chaotic fields in turn can lead to anomalous effects like plasma heating, particle acceleration and diffusion. These anomalous effects are essential to properly interpret some of the intriguing observations, e.g., solar coronal heating, energetic heavy ions in the vicinity of comets Halley and Giacobini-Zinner observed by recent cometary space missions.<sup>4,5</sup> Before proceeding with the mathematical analysis of our problem, we would like to ask ourselves the basic question: what exactly do we mean by chaos? There is no unique definition of chaos. The simplest

way it can be defined is as follows: If we map the trajectory of a particle in a given field and if the trajectory is smooth then the system is not chaotic. On the other hand, if the mapping is scattered, the system may be chaotic. Alternatively, a chaotic system, by definition, is very sensitive to the initial conditions e.g., an extremely minute difference in the initial conditions, of the particle, may lead two neighbouring trajectories to highly diverging trajectories. The necessary condition for the system to be chaotic is the exponential divergence of the two neighbouring trajectories.

### SOLITARY WAVES

Let us consider a plasma which is embedded in a magnetic field  $\vec{B}_0$  in the x-y plane. The electromagnetic waves propagating along x-direction in this plasma are governed by the two-fluid equations<sup>6,7</sup>. By means of singular perturbation method, we can show that these fluid equations alongwith the generalised Ohm's law, lead to the following *evolution equation* for the nonlinear Alfvén waves<sup>6,7</sup>:

$$\frac{\partial \vec{b}_\perp}{\partial t} + \alpha \frac{\partial}{\partial x} (\vec{b}_\perp | \vec{b}_\perp |^2) + \mu \left( \hat{e}_x \times \frac{\partial^2 \vec{b}_\perp}{\partial x^2} \right) = 0 \quad \dots(1)$$

where  $\vec{b}_\perp = (b_y, b_z)$ ,  $\mu = V_A / (2\Omega_i)$ ,  $V_A = B_0^2 / 4\pi\rho =$  Alfvén speed,  $\Omega_i = eB_0/m_i c =$  ion cyclotron frequency,  $\alpha = 1/4 (1-\beta)$  with  $\beta$  as the plasma  $\beta$  i.e., the ratio of the kinetic pressure to the magnetic pressure. Eq. (1) governs the elliptically polarized Alfvén waves; for circularly polarized waves this reduces to,

$$\frac{\partial \vec{b}_\pm}{\partial t} + \alpha \frac{\partial}{\partial x} (b_\pm | b_\pm |^2) \pm i\mu \frac{\partial^2 b_\pm}{\partial x^2} = 0 \quad \dots(2)$$

where  $b_\pm = b_y + ib_z$ ;  $b_+$  for left hand polarized mode and  $b_-$  for right hand mode. Throughout this paper we have used the normalised units<sup>6</sup>. We may note that Eq. (1) is a coupled nonlinear equation whereas Eq.(2) shows that left-hand and right-hand modes get decoupled. Eq. (2) is the Derivative Nonlinear Schrodinger (DNLS) equation, which can be solved exactly; its solution is given by<sup>8</sup>,

$$|b_\pm|^2 = \pm \frac{8(1-\beta)(V-V_s)}{\sqrt{2} \cosh[2(V-V_s)(x-V_s t)] \pm 1} \quad \dots(3)$$

where  $V$  is the phase velocity of the linear Alfvén wave ( $=V_A$  in unnormalized units) and  $V_s$  is the speed of solitary Alfvén wave. Eq. (3) is the solitary wave solution for left hand polarized Alfvén wave. Since the left hand side of (3) is positive definite, the solution with (+) sign will correspond to

$$(1-\beta) (V-V_s) > 0$$

and the one with (−) sign to

$$(1-\beta) (V-V_s) < 0$$

We may point out that the plasma  $\beta$  plays a very crucial role. For  $\beta < 1$ , super Alfvénic solitons have larger amplitude compared to the amplitude of the sub-Alfvénic solitons, but for  $\beta > 1$ , the behaviour is reserved. It is interesting to note that for laboratory plasmas,  $\beta < 1$  but for space plasmas,  $\beta > 1$  for many systems.

So far, we have considered only uniform plasmas. However, very often, we encounter non uniform plasmas. Repeating the procedure outlined above, we have derived the evolution equation for plasmas with weak but arbitrary inhomogeneity. The equation in this case is modified DNLS which by means of a series of complicated transformations can be reduced to DNLS<sup>9,10</sup>; once again we get solitary waves, which no more move with constant speed. These solitary waves get accelerated or decelerated depending on whether the wave propagation is along the increasing density gradients or decreasing density gradients. The evolution equation as well as its solutions are complicated, so they are not presented here (cf. ref. 9 and 10).

### CHAOTIC ALFVEN WAVES

Unlike equation (2), Eq. (1) cannot be solved analytically. For any further analysis of Eq.(2) and for investigating the possibility of chaotic behaviour of the system governed by this equation, it is useful to use the *Hamiltonian formulation*.

In a stationary frame of reference, namely,

$$\xi = (x - Vt)/\mu,$$

Eq (1) can be rewritten as

$$\hat{e}_x \times \frac{d\vec{b}_\perp}{d\xi} = \frac{d\Psi}{d\vec{b}_\perp} \quad \dots(4)$$

where

$$\Psi(\vec{b}) = \alpha \left[ \frac{1}{4} (b^2 - b_0^2)^2 - \frac{\lambda b_0^2}{2} (b - \vec{b}_0)^2 \right]$$

with  $b_0$  and  $\lambda$  as constants<sup>6,11,12</sup>. It is instructive to note that Eq.(4) is equivalent to a classical equation of motion of a particle, with zero mass, which is moving under the influence of the coriolis force and a pseudo potential  $\Psi$ . The equipotential contours of  $\Psi$ , thus represent the trajectories of this pseudo particle with zero mass. In terms of the Hamiltonian (H) which in this case is nothing but  $\Psi$  Eq(4) can be represented by a set of following two equations.

$$\frac{db_y}{d\xi} = \frac{\partial H}{\partial b_z} \quad . \quad (5a)$$

$$\frac{db_z}{d\xi} = - \frac{\partial H}{\partial b_y} \quad (5b)$$

We immediately recognize these equations as Hamilton's equations with  $b_y$  and  $b_z$  as canonical co-ordinates. The potential  $\Psi$  being a quartic has four solutions, namely, the minimum at  $(\lambda - b_0)$ , the maximum at  $(\lambda - b_0)$  and the saddle point at  $b_0$ . For  $0 < \lambda < 1/2$ ,  $\lambda_\pm$  are given by<sup>11</sup>,

$$\lambda_\pm \equiv \frac{1}{2} [-1 \pm (1 + 4\lambda)^{1/2}] \quad \dots(6)$$

For this range of  $\lambda$ , the equipotential contours of  $\Psi$  and the three solutions corresponding to dark, bright and mixed solutions are shown in Fig. 1. For  $\beta < 1$ , dark (bright) solitons have right (left) hand polarization. These solitons do not interact with each other and hence there is absolutely no possibility of having a chaos.

## DRIVEN HAMILTONIAN SYSTEMS

From Fig. 1, it is very apparent that the solitary (localized stationary nonlinear) Alfvén waves, left to themselves, will propagate as such for ever but this is somewhat an ideal situation. We would like to find out what happens to these solitary waves when there is some local disturbance e.g., if there is a possibility of having another plasma wave generation or there is

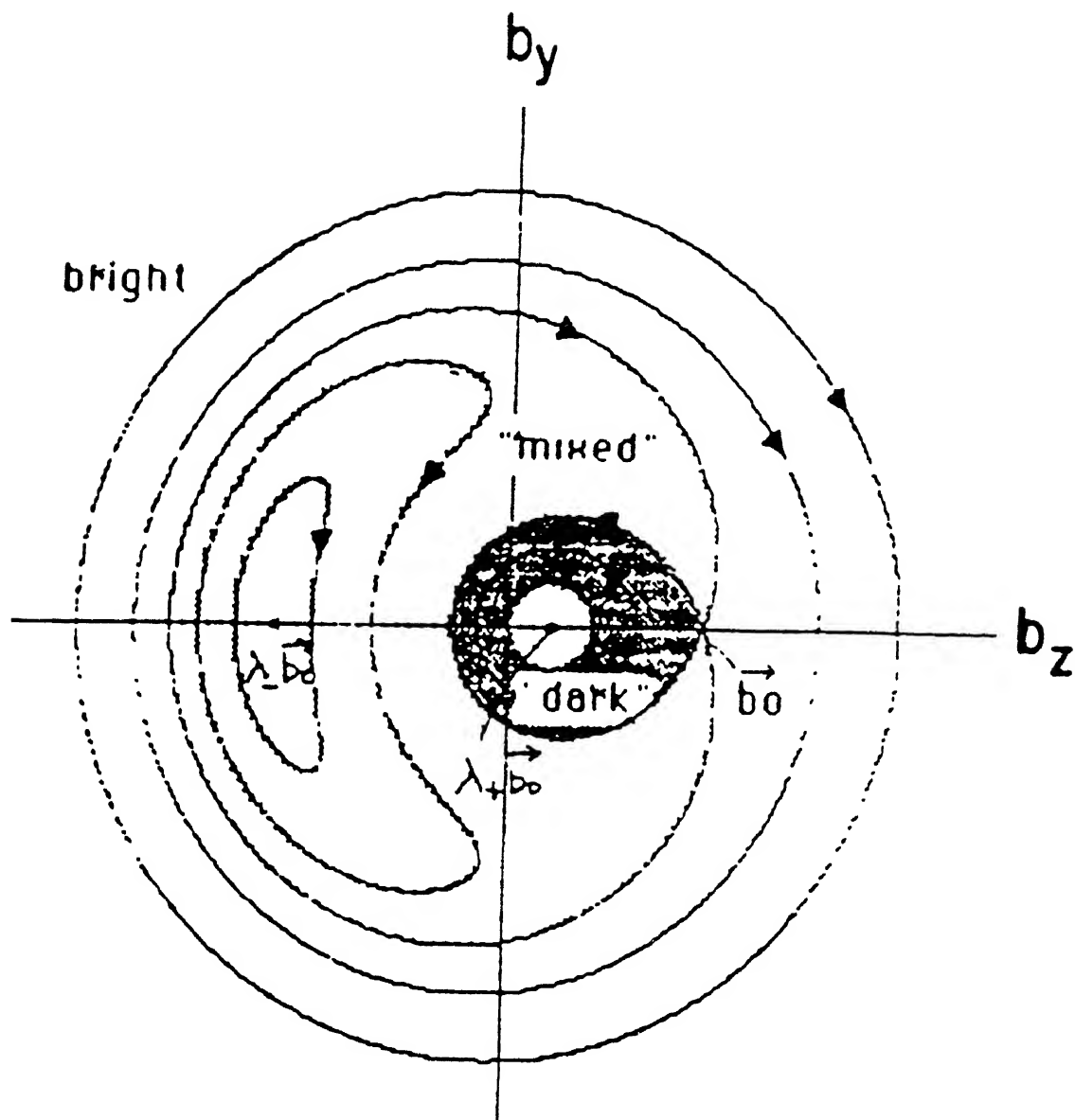


FIG. 1 Equipotential contours of the pseudo potential  $\psi$ . also shown are 3 different solutions namely dark, bright and mixed solitons. Mixed solitons correspond to  $\psi < 0$ .

some solar activity (say solar flare) which could propagate and interact with the Alfvén waves. To explore this, we will first consider the external driver as a plane wave. In order to make the problem tractable i.e., to be able to reduce the coupled partial differential equations (given by Eq.(1)), modified by the driver, to a set of ordinary differential equations like Eq. (5), we have to further restrict ourselves to the driver which is stationary in the frame of reference of the stationary Alfvén waves. This is indeed a very big

restriction; in the latter part of this lecture we will remove this restriction. Eq.(5), in the presence of such a driver, is replaced by<sup>6,7</sup>,

$$\frac{db_y}{d\xi} = \frac{\partial H}{\partial b_z} + A \cos \theta \quad (7a)$$

$$\frac{db_z}{d\xi} = -\frac{\partial H}{\partial b_y} + A \cos \theta \quad (7b)$$

$$\frac{\partial \theta}{\partial \xi} = \Omega \quad (7c)$$

In Eqs.(7),  $A$  is the amplitude of the driver and  $\Omega$  is its frequency. We will now show that for  $A \neq 0$ , the dark, bright and the mixed solitons start interacting among themselves and when the driver is sufficiently strong i.e.,  $A$  is sufficiently large, this interaction can lead to chaos. The reason for the appearance of chaos is rather transparent. In the absence of the driver, our system has two degrees of freedom (cf. Eqs.(5)) but in the presence of the driver, it has three degrees of freedom (cf. Eqs.(7)) and from our knowledge of nonlinear dynamics, we know that the minimum number of degrees of freedom required, for a system to be chaotic, is three and hence the possibility of chaos with the driver. For the left hand polarized driver i.e.,  $\Omega = -2$ , our numerical results are summarised by Poincare maps shown in Fig. 2. For  $A = 0$ , the entire set of Poincare points, originating from a given initial point, remains on the potential contour containing that initial point. For  $A = .002$  i.e., for a weak driver, one of the sets of Poincare points, near the bright soliton separatrix, starts to scatter in limited region of phase space. This leads to the onset of chaos for this set of Poincare points. We notice the formation of 3 islands near the local minimum of the potential  $\psi$ . The number 3 is not a magic number; this is simply because we have taken  $\omega/\Omega \sim 2/3$  ( $\omega$  being the frequency of Alfvén waves). For somewhat stronger driver e.g., for  $A = 0.2$ , the system is almost chaotic except for the region inside the dark soliton (cf. Fig. 1). This region is unaffected because of the difference in polarity of the driver ( $\Omega < 0$ ) and the Alfvén waves ( $\omega > 0$ ). There is little interaction of the driver and the Alfvén waves because of opposite polarities. We repeated these calculations with the right hand polarized driver ( $\Omega = 2$ ); these results are shown in Fig. 3. Comparison of Poincare maps for  $A = 0.2$  for  $\Omega < 0$  and  $\Omega > 0$  testifies our above drawn conclusion. No matter how strong is the

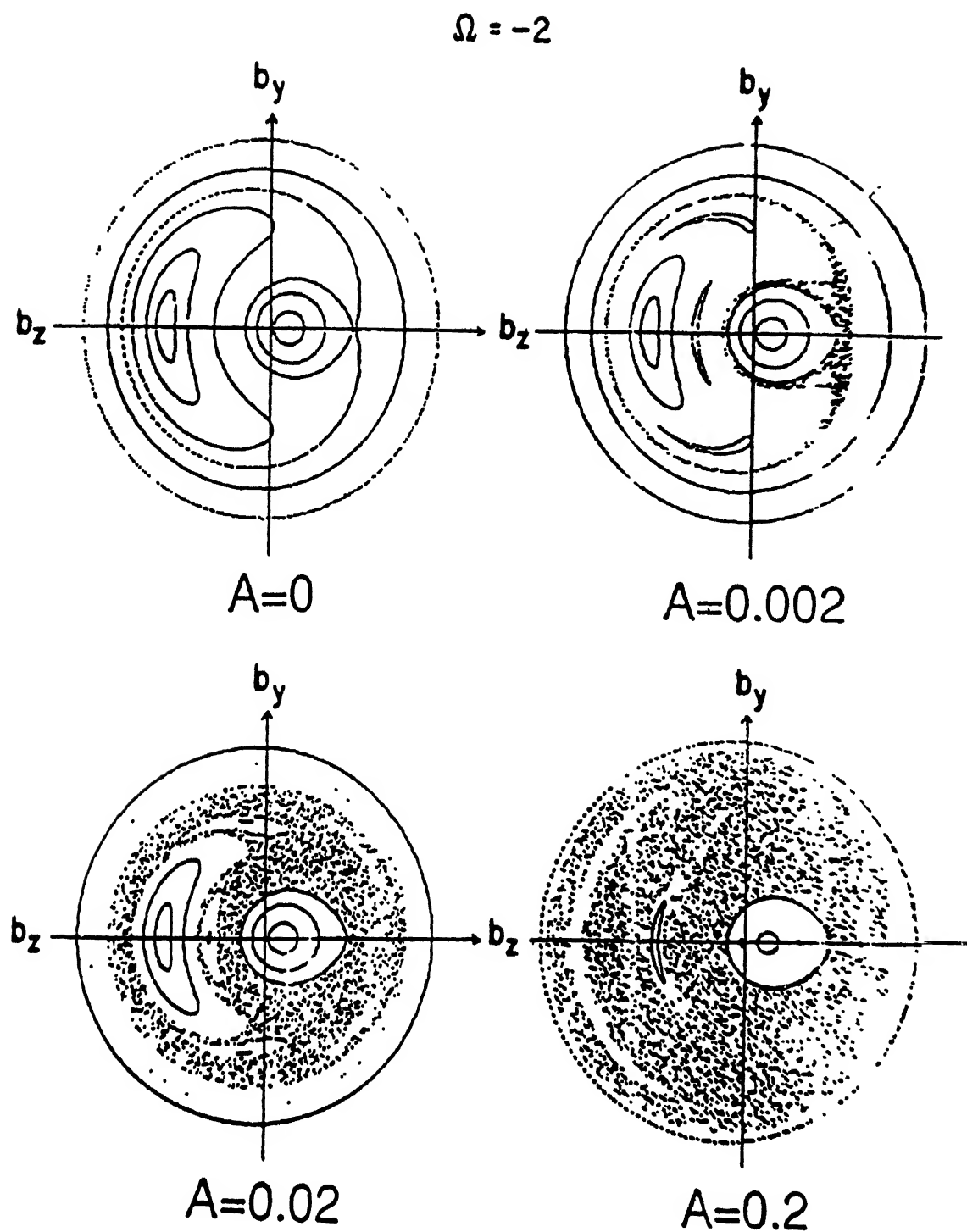


FIG. 2 Poincaré maps for the left hand polarized driver ( $\Omega = -2$ ) for various amplitudes of the driver.

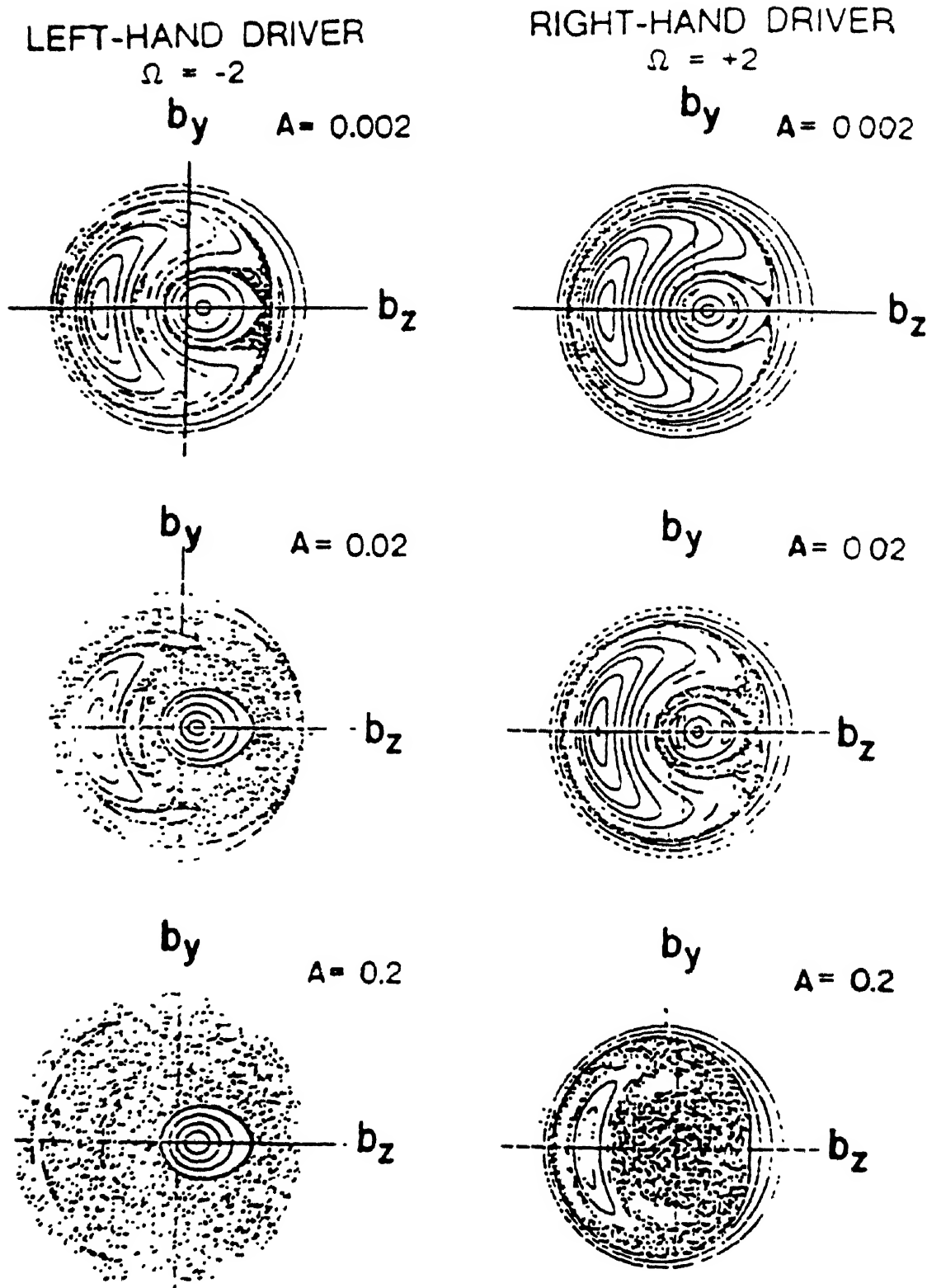


FIG. 3 Comparison of Poincare maps for the left and the right hand drivers. The left hand column is the repetition of Fig. 2.



driver, the region inside the dark soliton remains coherent for  $\Omega < 0$ . It is this what we call *Coherence in Chaos*

## MULTISPECIES PLASMAS

So far our discussion has been confined to plasmas with only two species e.g., hydrogen plasma. Very often, we encounter multispecies plasmas with electrons, protons and heavy ions; the heavy ions, in laboratory plasmas, may be as impurities and in some natural plasmas as a genuine constituent, e.g., solar wind is composed of electrons, protons and  $\alpha$ -particles (helium).

To study the chaotic processes in such multispecies plasmas, starting from the corresponding two fluid equations we derived the evolution equation<sup>8</sup> which is not shown here. For solar wind parameters with 5% helium in abundance, we repeated our calculations; Poincare maps, for the two-species and the three-species plasmas are shown in Figs. 4 and 5 respectively<sup>13</sup>. For both the figures, the driver is the left hand driver. Since all the parameters for both the figures are the same, Fig. 5 shows the effect of  $\alpha$  particles. From comparison of these two figures, it is evident that the chaos is reduced due to the presence of helium — in other words, the threshold for chaos goes up because of heavy ions. Physically this could be interpreted as the inertial stabilisation due to heavy ions.

## STRANGE ATTRACTORS

For Hamiltonian systems, it is well known that the phase space volume is conserved but for dissipative systems, the phase space ( $b_y - b_z$  in our case) volume undergoes continuous contraction with increasing time. This leads to phase space of lower dimensionality and eventually to what are known as attractors. For nonchaotic systems, the attractors are simple attractors e.g., a point, a line or a limit cycle whereas for chaotic systems, we can have attractors with very complicated structures, these are known as strange attractors.

Our entire discussion, so far, has been about the chaotic phenomena in nondissipative plasmas. Now we will very briefly present our efforts in this direction in connection with the dissipative systems. You notice that we had started with the simplest possible system and have been introducing the complexities into our system one by one. This step by step approach is simply to get a good physical insight into the complex phenomena of chaos.

For a dissipative two-species plasma, the nonlinear evolution equation for stationary Alfvén waves is given by<sup>6,12</sup>

## 2 - SPECIES

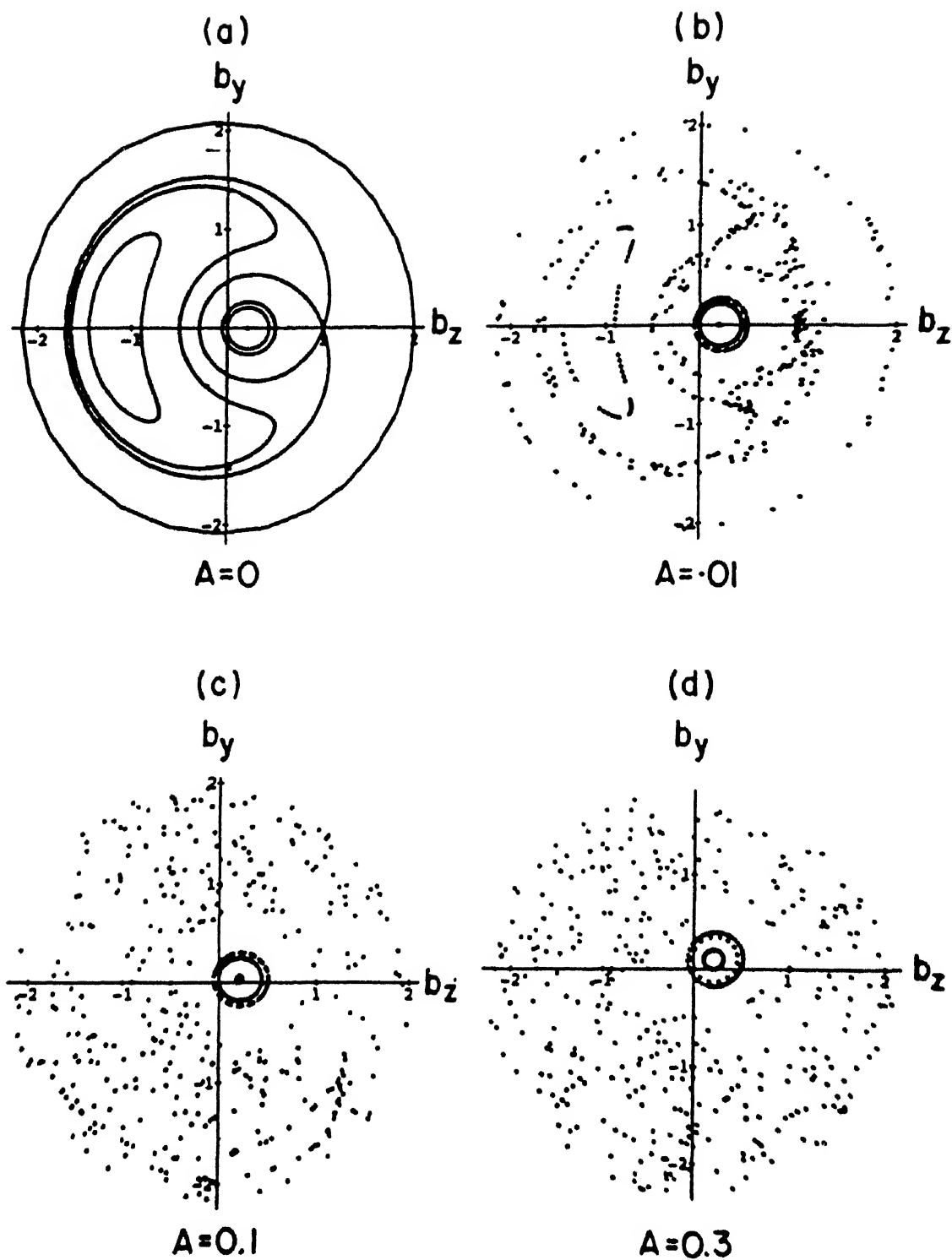


FIG. 4 Poincaré maps for driven Hamiltonian system for solar wind at 1 AU with only two species (electrons and protons). Driver in this case is left hand polarized.

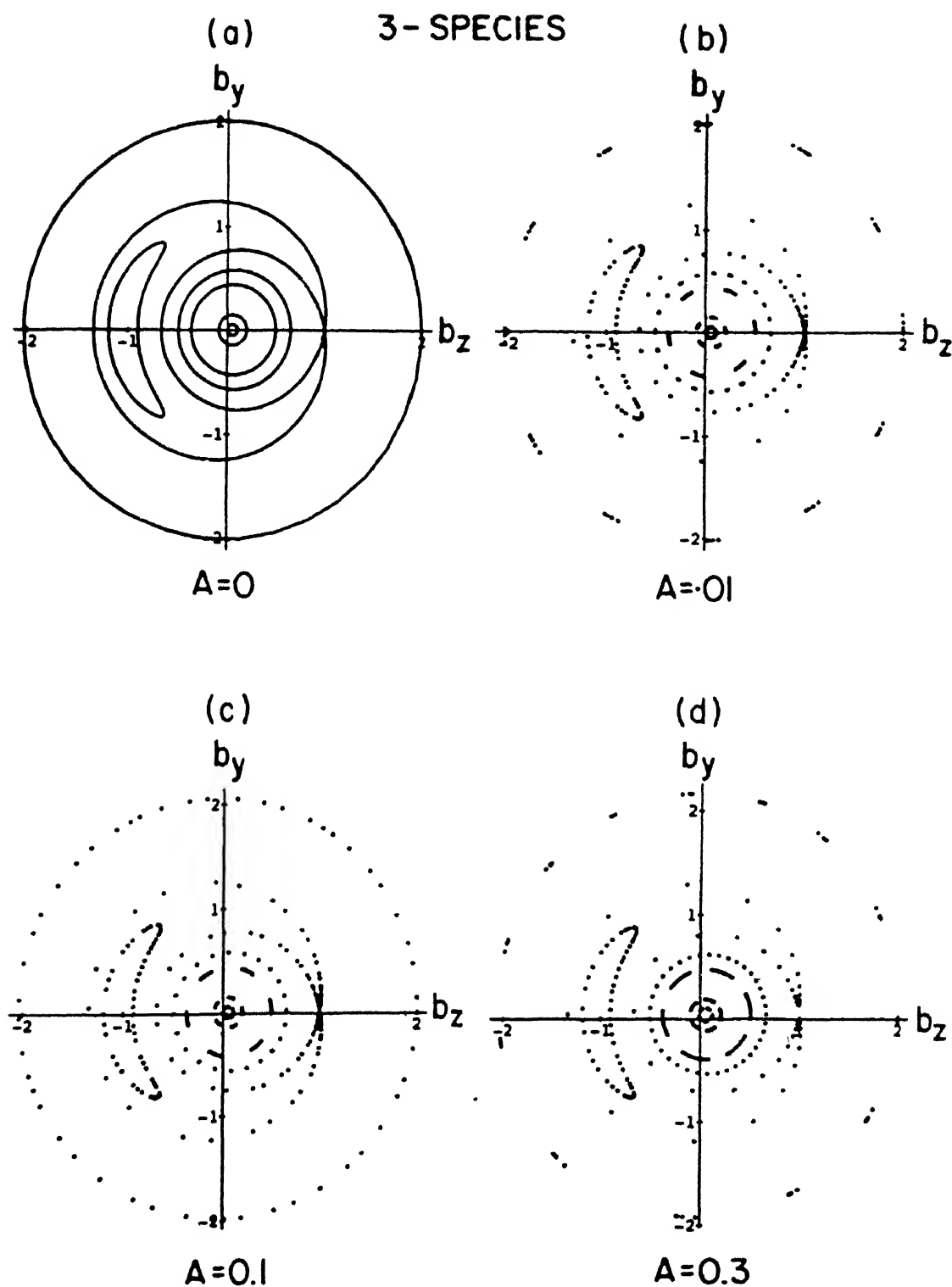


FIG 5. Same as Fig. 4 but for solar wind with three species i.e., electrons, protons and  $\alpha$  particles;  $N_\alpha/N_p = .05$ . This figure is valid also for cometary plasma with water group ions as the third species and with  $N_i/N_p = .01$  and  $m_i/m_p = 16.8$  ( $i$  represents heavy ions and  $p$  represents protons).

$$\hat{e}_x \times \frac{d\vec{b}_\perp}{d\xi} + \nu \frac{d\vec{b}_\perp}{d\xi} = \frac{\partial \Psi}{\partial \vec{b}_\perp} + A \begin{pmatrix} \cos \Omega \xi \\ \sin \Omega \xi \end{pmatrix} \quad \dots(8)$$

where  $\nu > 0$  for dissipation and in the last term on right side of Eq.(8), we have  $\cos(\Omega \xi)$  for  $b_y$  and  $\sin(\Omega \xi)$  for  $b_z$  component of the vector equation.

Following the procedure outlined in the preceding sections, we have done the detailed analysis of Eq.(8) and have come up with some very interesting results. Summary of these results is presented in Figs. 6 and 7. Fig. 6 shows the orbits as well as the Poincare points (shown by dots) in the  $b_y - b_z$  phase space. The number of dots represents the periodicity. For  $A = 0$ , we found two attractors: one at the minimum of the potential and the other one at the saddle point. Figs 6 and 7 show the evolution of these attractors as  $A$  increases. The most interesting observation that we make is the complicated structure of the attractor for  $A = 0.16$  (shown in I). In this box, we have not drawn the orbits but only the Poincare points are shown. This attractor is a strange attractor; this we have confirmed by checking its self similar property and by determining its fractal dimension which is found to be 1.57. For further discussion of Fig. 6, refer to Hada et al<sup>12</sup>. Another very interesting observation we make from Fig. 7 is that the chaos arises through two very distinct channels. At the saddle point, there are three sporadic attractors at  $A \sim 0.035$ ,  $A \sim 0.042$  and  $A \sim 0.102$ . The chaos in this channel arises through a sequence of period doubling bifurcations. However, at the minimum of the potential, in the region of strange attractor, there is a sudden transition to chaos. For more details, see Hada et al<sup>12</sup>.

## CONCLUSIONS

The techniques, used in this paper for the study of chaotic Alfvén waves, can be used for the study of chaotic phenomena in any other system. Here we have discussed only the stationary waves with a very special kind of driver. However, this restriction on the driver can be relaxed; we can consider any other driver e.g., a beam representing solar wind or a pulse representing solar burst or a wave packet or any other type but then we cannot reduce our partial differential equation (cf. Eq. (1)) to ordinary differential equations and we have to solve partial differential equations. We now have a code for solving PDE's; this solution gives us the time series (time evolution) for  $b_y$  and  $b_z$ . From the time series, like in the case of observed time series, we have calculated the energy spectrum, correlation

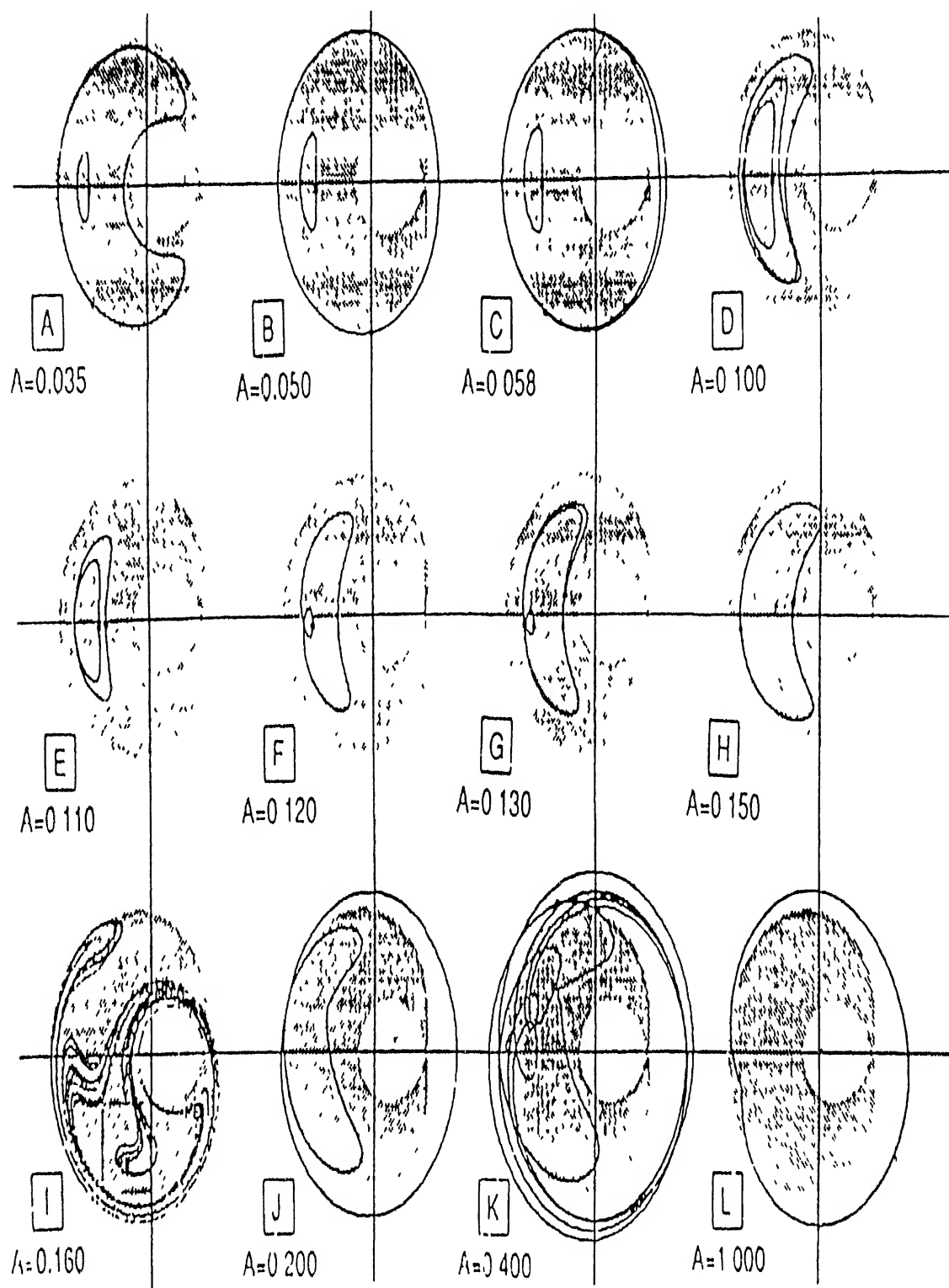


FIG. 6 Shows attractors for the driven dissipative system in  $b_y - b_z$  phase space. The labels are the same as in Fig. 7. The dots on the trajectories are the Poincaré points. In (I), trajectories are omitted and only the Poincaré points are plotted. The complicated structure of the attractor in this case represent a strange attractor for  $A = 0.16$ .

## Bifurcation Diagram

$$\Omega = -1, \nu = 0.02$$

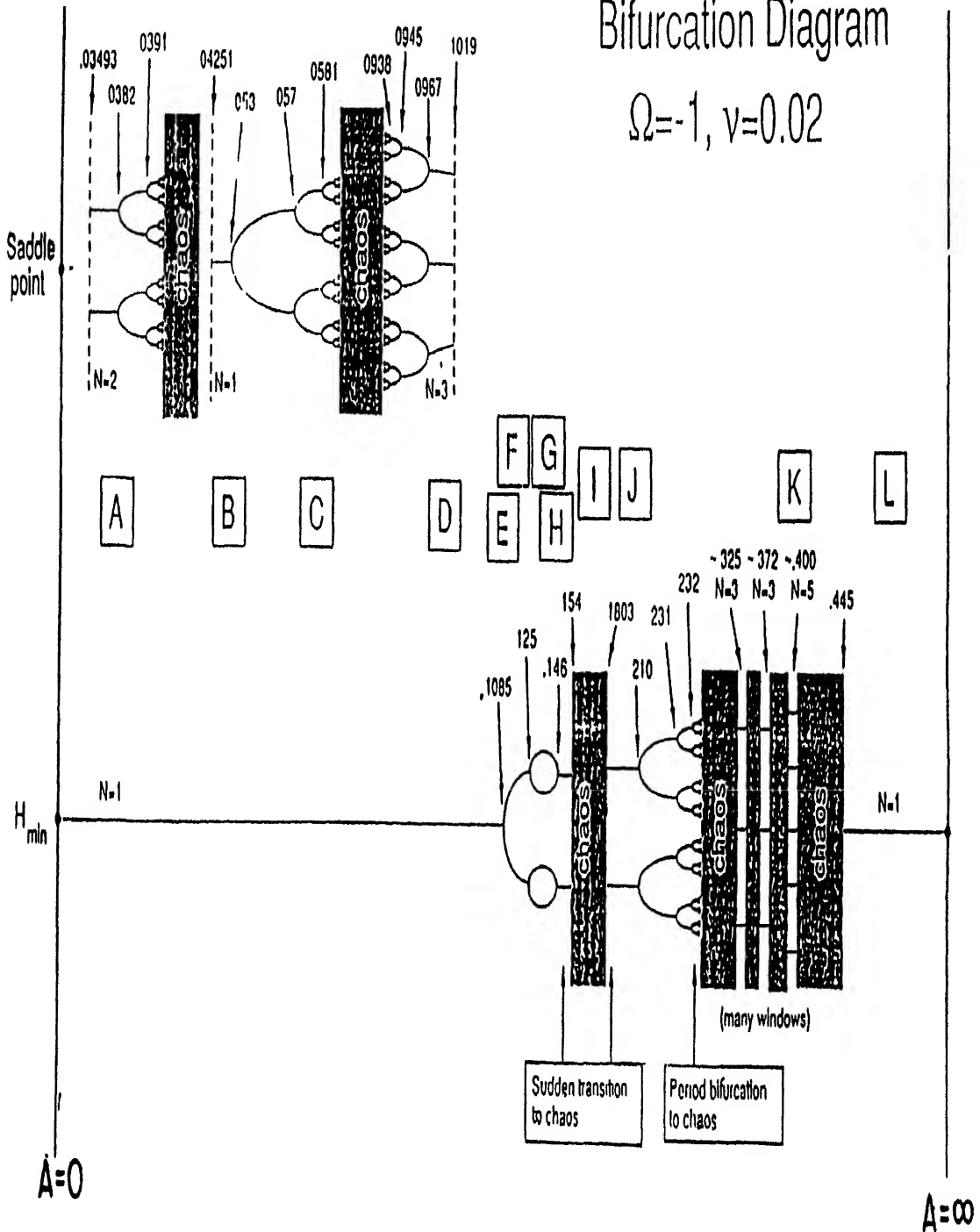


Fig. 7 Shows the evolution of various attractors as the amplitude of the driver increases. The numbers on the top side correspond to the amplitudes. The top channel shows the normal as well as the reverse bifurcations. The bottom channel gives sudden transition to chaos in the neighbourhood of the strange attractor at  $A = 0.16$ .

dimension and Lyapunov exponents which characterise the phenomena of chaos. The latter calculations for a variety of drivers are under progress. The different spectra obtained for different drivers will be compared with the observed spectrum to finally conclude about the source of observed Alfvénic turbulence. As mentioned earlier, here we have considered Alfvén waves but the procedures outlined here are applicable to any other mode.

Chaotic systems can be very efficient sources of particle acceleration and plasma heating<sup>14,15</sup>. We plan to investigate the anomalous acceleration of heavy ions observed in the vicinity of comets and anomalous solar coronal heating due to chaotic magnetic fields discussed in the present paper.

I would like to express my gratitude to all my collaborators for the contributions presented in this lecture.

## References

- 1 F L Scarf, F V Coroniti, C F Kennel, D A Gurnett and E J Smith, *Science* **232** (1986) 382
- 2 L F Burlaga, *Rev. Geophys Space Phys* **21** (1983) 363
- 3 B T Tsurutani and E J Smith, *Geophys Res Lett* **13** (1986) 259.
- 4 S McKenna-Lawlor, E Kirsch, D O'Sullivan, A Thompson and K P Wentzel, *Nature* **321** (1986) 299
- 5 R J Hynds S W H Cowley, T R Sanderson, K P Wentzel and J J Rooijen, *Science* **232** (1986) 361
- 6 B Buti, *Solar and Planetary Plasma Physics*, Ed B Buti, *World Scientific, Singapore* (1990) 92
- 7 C F Kennel, B Buti, T Hada and R Pellat, *Phys Fluid* **31** (1988) 1949
- 8 F Verheest and B Buti, *J Plasma Phys* **47** (1992) 15
- 9 B. Buti, *Geophys Res Lett* **18** (1991) 809
- 10 B Buti, *J Plasma Phys* **47** (1992) 39.
- 11 T Hada, C F Kennel and B Buti, *J Geophys Res* **94** (1989) 65.
- 12 T Hada, C F Kennel, B Buti and E Mjølhus, *Phys Fluids* **B2** (1990) 2581
- 13 B Buti, *Jou Geophys Res* **97** (1992) 4229
- 14 B Buti and G S Lakhina, *Geophys, Res Lett* **14** (1987) 107.
- 15 B Buti, *Cometary and Solar Plasma Physics*, Ed. B Buti, *World Scientific, Singapore* (1988) 221



**Kazhiyur Kothandapani Kannan** (b. 30 September 1939) did Ph.D (1966), Indian Institute of Science, Bangalore. He is Scientific Officer (G), Solid State Physics Division, BARC, Bombay.

Kannan has been engaged in X - ray diffraction analysis of single crystals. He has worked on biological macromolecules like proteins and viruses and uncovered two protein structures, those of human carbonic anhydrase I and II, their inhibitor and drug complexes and also of satellite tobacco necrosis virus - the first protein and virus structures from Sweden. Proposed the functional mechanism of the enzyme by pin-pointing the important active site residues of the proteins, which have been confirmed by site - directed mutagenesis experiments. He and his colleagues are the first in India to have isolated, purified, crystallized and solved the protein structure of buffalo carbonic anhydrase II. He has used synchrotron radiation for protein structure - function - drug interaction work. Crystallized a multienzyme complex containing Rubisco from spinach. The structure is being solved by molecular replacement method using the synchrotron X- ray diffraction data. He has also started work on directed mutagenesis of human carbonic anhydrase I to elucidate the function and folding of the enzyme. He has contributed chapters to 'The Enzymes' and 'Advanced Methods in Protein Sequence Determination'.

Kannan is Fellow of Indian Academy of Science; Life Member, Indian Physics Association and Indian Biophysical Society; he is Secretary INSA (1993-95). He is recipient of Jawaharlal Nehru Birth Centenary Lecture Award (INSA) (1992).

*Kazhiyur Kothandapani Kannan was elected to the fellowship of the Academy in 1987.*

Address: Scientific Officer (G), Solid State Physics Division, Bhabha Atomic Research Centre, Bombay 400 094; Tel: (022) 5563060 (Ext 2152); Fax: 91-22-5560750; E-mail: kannan@magnum.barcl.ernet.in, Res: 17-C, Nandadevi, Anushakti Nagar, Bombay 400 094; Tel: (022) 5567848.



# **INSIGHT INTO FUNCTION OF AN ENZYME FROM THE MOLECULAR ARCHITECTURE**

K K KANNAN FNA

Life is a complex process made possible by the innumerable interactions of ions, small molecules and macromolecules like proteins. Indeed the vital life functions are made possible by a large number of enzymes in any given organism. It is therefore of paramount importance to understand how proteins are made up of and how the molecular architecture i.e. the structure of any protein is intimately connected with its function. This in turn requires the elucidation of the three dimensional atomic structure of such complex molecules as proteins and viruses and their interaction with other small molecules such as drugs and large molecules which either make them work or stop them from functioning. I will trace in this lecture some rudiments of protein structure determination and illustrate with the work on carbonic anhydrase isoenzymes which I and my colleagues Drs. M.Ramanadham, V.S. Yadava, Vinay Kumar, Shri. S. Chakravarty and Smt. Padma Sathyamurthy have done at BARC, Bombay. We have other projects like the RUBISCO - multienzyme complex which is being carried out by me and my colleague Dr. M.V.Hosur in collaboration with Dr. J. Saini of the Molecular Biology and Agriculture Division of BARC. In this lecture I will confine myself only to carbonic anhydrases which have been investigated by us and others over a much longer time period yielding very important information on the structure and function of these isozymes.

## **PROTEIN CRYSTALLOGRAPHY**

The advent of Protein Crystallography can be traced to the first X-ray diffraction pictures of protein and virus crystals, by Bernal, Crowfoot (now Dorothy Hodgkin) in 1934. By then X-ray diffraction studies of organic and inorganic crystals were becoming routine even though there was a lack of sophistication in the techniques used. The blossoming of protein crystallography is of course solely due to the structure determination of Myoglobin by Kendrew and coworkers and Haemoglobin by Perutz and coworkers. Perutz established protein crystallography as a formidable science when he and Ingram showed the efficiency of isomorphous

replacement method (MIR) using heavy atom modifications of proteins in crystals.

It is obvious that protein crystallography requires protein crystals and that too good ones and in large numbers for the different investigations. Getting protein crystals in quality and quantity is a less well understood science and has been mainly based on trial and error. The pre-requisite for such experiments is well purified and characterized proteins of interest in fairly good quantity of the order of 10s of milligrams. Ammonium sulphate, polyethylene glycol (PEG), Methane Pentane Diol (MPD) are some of the solvents of choice for crystallization trials. The salt concentrations, *pH* and other additives are varied in micro droplets in the experiments in screening for crystals.

Once crystals have been obtained they are then characterized by x-ray diffraction methods. The packing of the protein molecules in the smallest repeating unit of the crystal is known as the unit cell. The dimensions of this unit cell in terms of three axial lengths and three angles between these axes are to be determined first. These can be done from photographs of x-ray diffraction known as precession pictures or automatically in the advanced

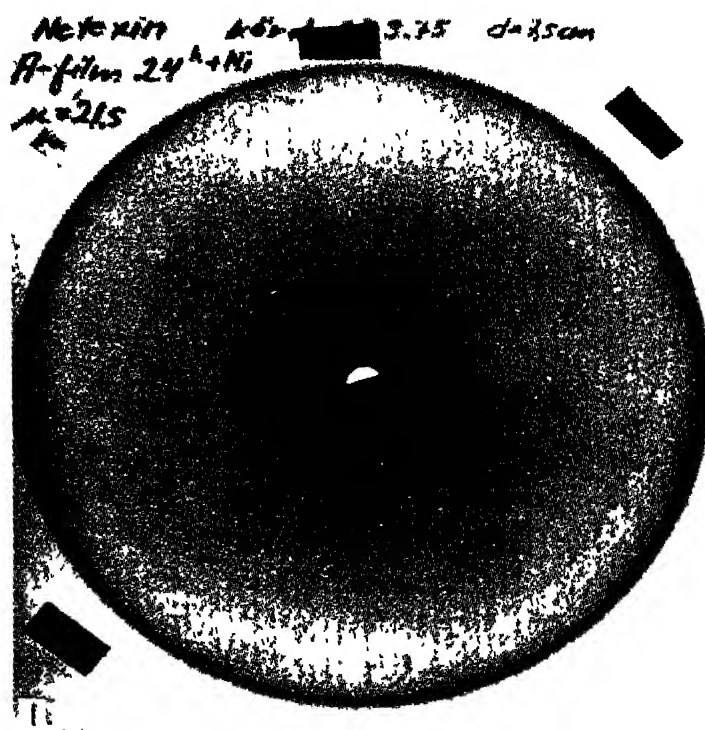


FIG 1 An x-ray diffraction photograph.

area detector diffractometer instruments available today. Conceptually the photographic film illustrates an important aspect of crystallography. The coordinates  $(h,k,l)$  of the dark spots called reflections contain all the relevant information regarding the atomic arrangement in a crystal. Thus no two crystals may have identical diffraction pattern unless the crystal is composed of the same compound and arranged spatially in the same way. Though the intensities of the reflection and the  $(hkl)$  coordinates contain all the necessary information about the molecule of interest packed in the crystal, it is not possible to obtain structural information directly from x-ray diffraction studies, as one does with light or electron microscope. The reason for this is the relationship between the intensities that are observations and the structure factor which is required to determine the structure. This is expressed as

$$|F(hkl)|^2 \propto I(hkl),$$

where  $|F(hkl)|$  is the amplitude of the reflection and  $I(hkl)$  is the intensity. The electron density distribution in a crystal is given by

$$\rho(xyz) = (1/v) \sum_h \sum_k \sum_l |F(hkl)| \exp[-i\{2\pi(hx+ky+lz) - \alpha(hkl)\}]$$

where  $\alpha(hkl)$  is known as the phase angle and  $i = \sqrt{-1}$  an imaginary number.

$$F(hkl) = |F(hkl)| \exp i \alpha(hkl) = \sum_{i=1}^n f_i \exp 2\pi(hx + ky + lz) = A + iB$$

$$\text{and } \alpha = \tan^{-1} (B/A)$$

Unless therefore the  $\alpha(hkl)$ s are determined the information contained in  $I(hkl)$  and  $(hkl)$  cannot be utilized to determine the structure. This is known as the "phase problem" in crystallography. Crystal structure determination of a crystal containing either small molecules or large molecules or a complex of these, *de facto*, requires the 'phase problem' to be solved first. This is where the work of Perutz in devising the multiple heavy atom isomorphous replacement method (MIR) for protein structure determination has played such a crucial role in protein crystallography.

Once a protein has been purified, crystallized and characterised initially, the intensity data  $I(hkl)$  is recorded for further use. Heavy atom

substitution in the protein crystals are attempted by modifying potential side chain groups such as cystine, histidine etc. or by heavy atom substituted inhibitor molecules which bind well to the protein. Sometimes prosthetic groups such as metal ions bound to the protein can be exchanged for heavier ones and profitably used in the MIR method. The heavy atom derivatives so produced are then characterized and the position of the bound heavy atoms determined by the difference Patterson method.

$$P(uvw) = |\Delta F(hkl)|^2 \cos 2\pi(hu + kv + lw)$$

This method does not require the knowledge of the phase information and gives the vectors ( $uvw$ ) between heavy atoms located in the crystal.

It is not always necessary to make heavy atom derivatives for successfully determining a protein structure. Over 300 protein structures are known today and an analysis of these structures has shown the similarities in protein structures and that proteins can be classified into a number of subgroups. Besides the structure of a protein which is homologous to the one of interest at hand may be expected to have similar three dimensional structure. The atomic coordinates of this may be available in the protein data bank (PDB). In such a situation one can make use of this information and try and solve the new protein structure of interest. A method of structure determination has been developed by Rossman and Blow for such a situation and has been exploited by many practising protein crystallographers successfully in the elucidation of many protein and virus structures. This is known as the "molecular replacement" method. The mathematical operation relating the orientation and position of the known molecule with respect to the unknown molecule in its unit cell can be expressed as

$$\vec{X}' = [C]\vec{X} + \vec{D},$$

where  $X$  is the coordinates of the known structure,  $X'$  that of the unknown structure,  $[C]$  is a rotation matrix and  $D$  is a translation vector. The concept behind this method is simple. One superposes the Patterson functions of the unknown structure over that of the known structure and rotate the known over the unknown and calculate the sum product of the two superposed Patterson functions. When the Patterson functions match with each other during this process the value for this product would be maximum and would give us the angle of rotation of the known protein. when it

would be similarly oriented as the unknown one. This is known as the "Rotation Function (RF)". Once the rotational angle is known then the position of the unknown protein molecule in the crystal in which it is present should be determined by the "Translation Function (TF)" technique. Here again only Patterson functions are used, except that the properly oriented known protein molecule is matched with the Patterson function of the symmetry related molecule in the crystal of the unknown but similar protein.

$$\text{Thus } \vec{X}' = [c]\vec{x} + \vec{d}$$

where  $\vec{d}$  is the translation between the two crystallographically symmetry related molecules  $\vec{X}'$  and  $\vec{x}$  rotated by the matrix defined by  $[c]$ . This can for example be a pure rotation or a screw rotation.

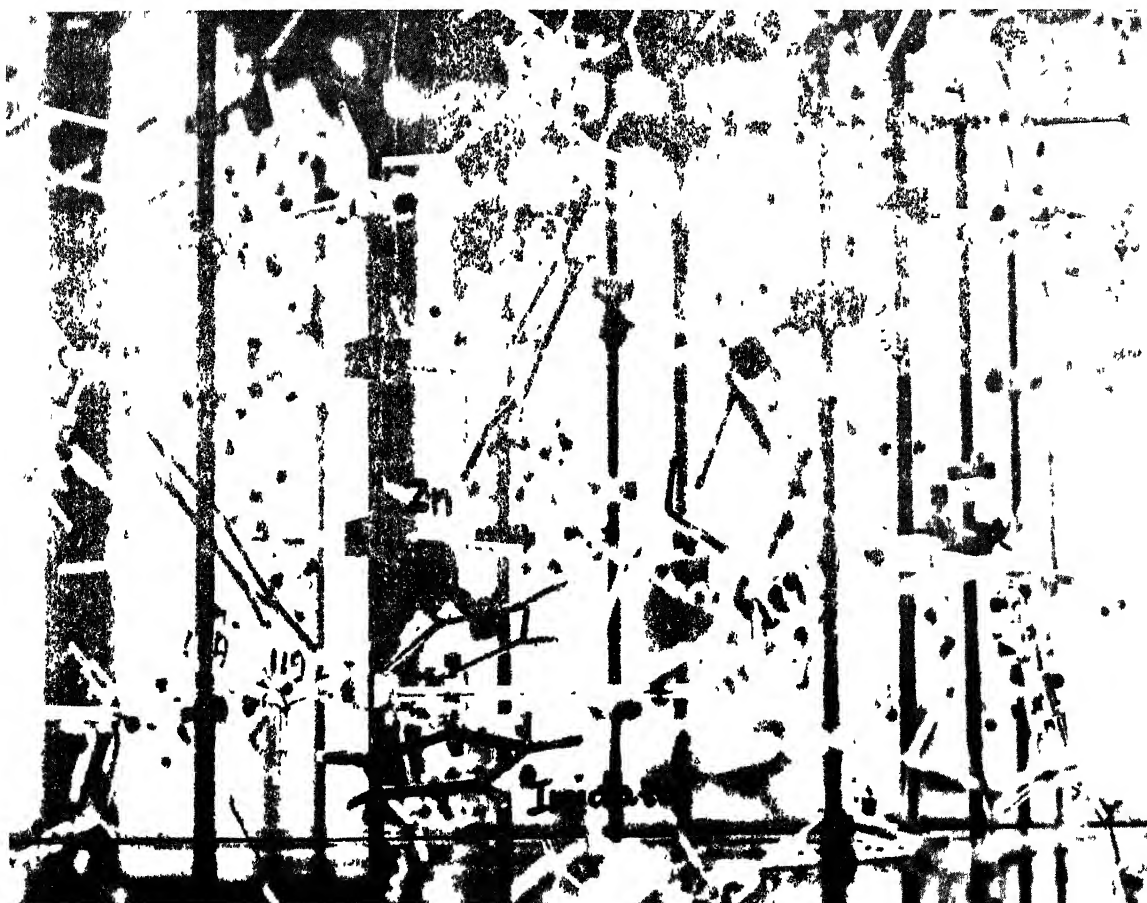


Fig 2 Kendrew Watson Skeletal Model parts built together.

Thus the rotation function and translation function can be used, and has also been used successfully, to determine the structure of a number of new protein structures when the structure of a homologous one is known. The success of rate for molecular replacement method is quite impressive, though some protein structures have not yielded to this method.

Once the appropriate phase angles  $\alpha(hkl)$  for a majority of the reflections are known from the MIR or MR method, an electron density can be calculated and the structure of the protein of interest can be interpreted from the electron density. The knowledge that shapes of amino acids should have correlation to the shape of the electron density is very handy in interpreting the electron density maps. In the early days of protein

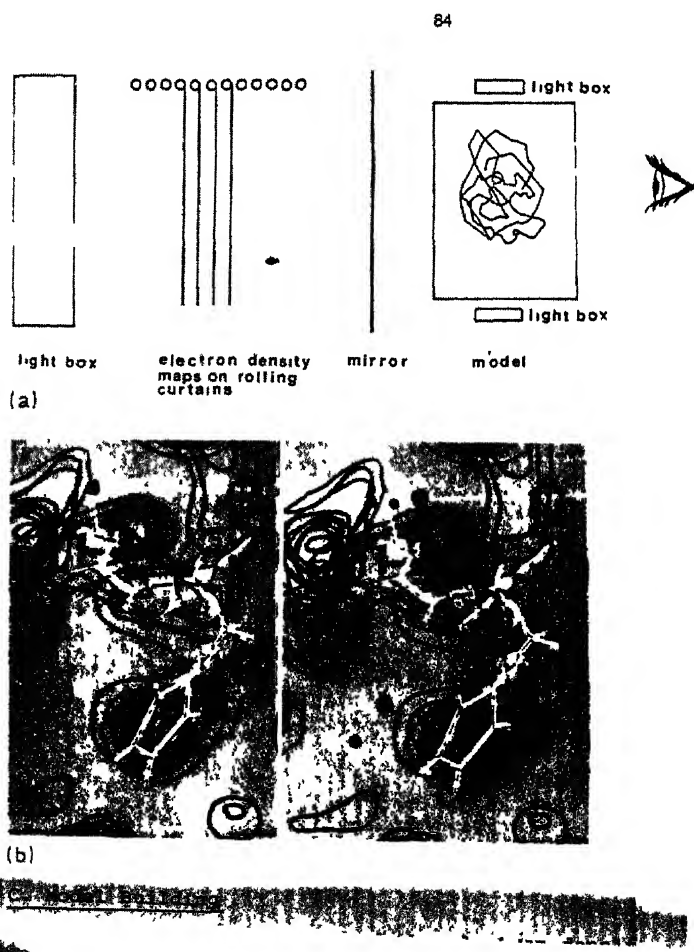


Fig 3 Schematic of Richard's Optical comparator showing histidine matched in the electron density.

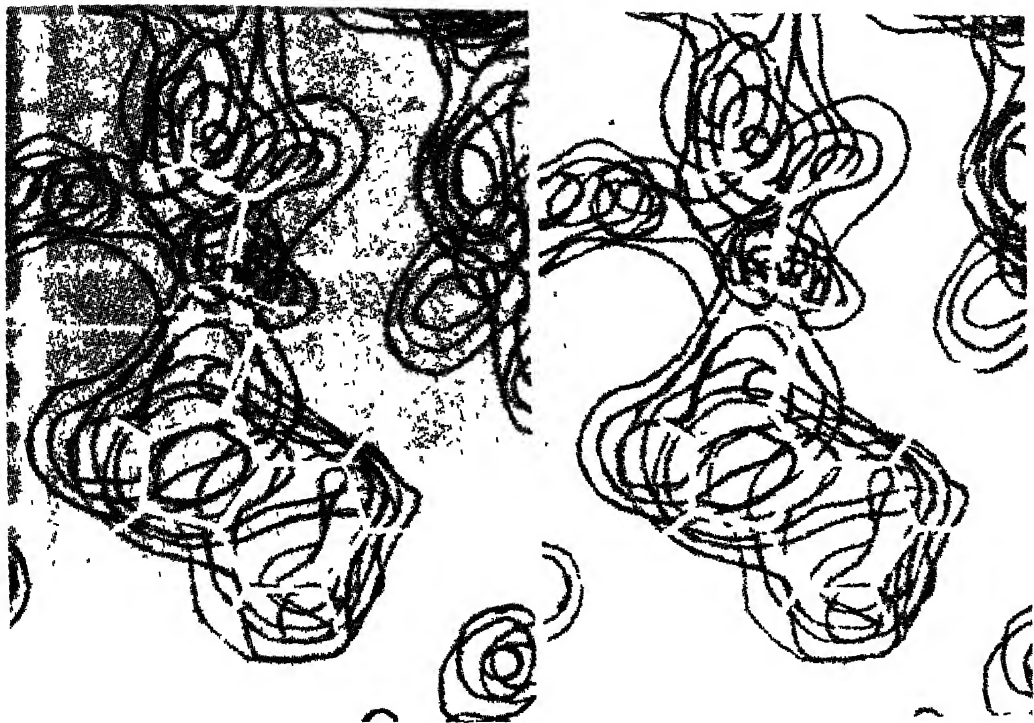


Fig 4a Tryptophane superposed in its electron density map in the optical comparator

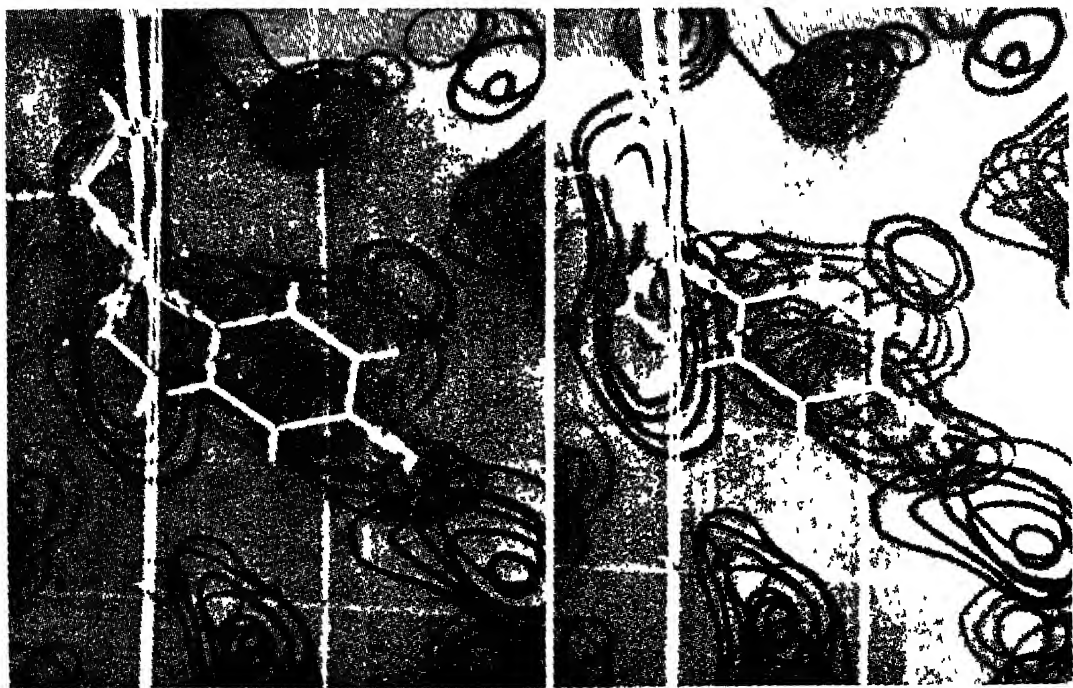


Fig 4b Tyrosine superposed in its electron density map in the optical comparator.

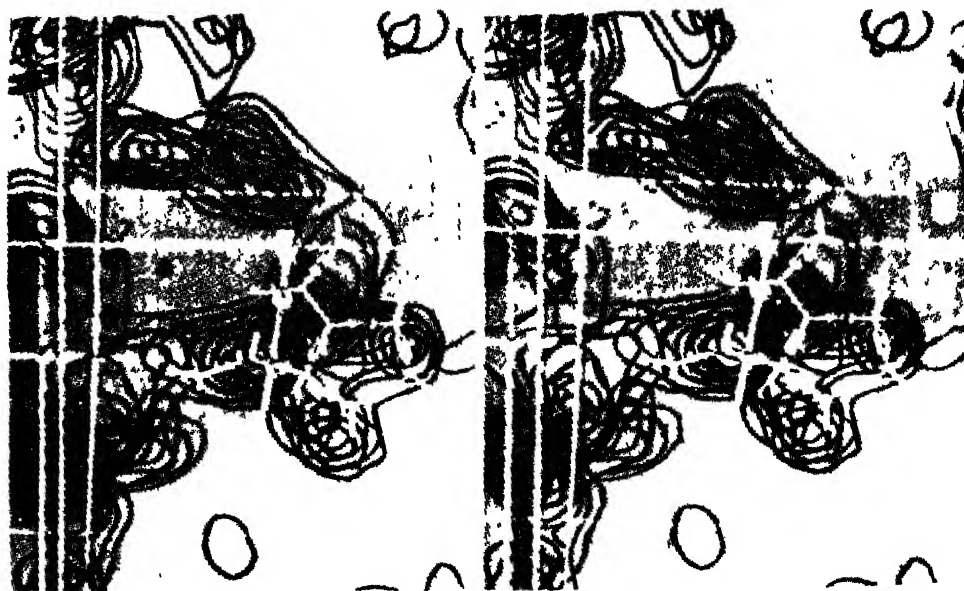


Fig 4c Proline superposed in its electron density map in the optical comparator

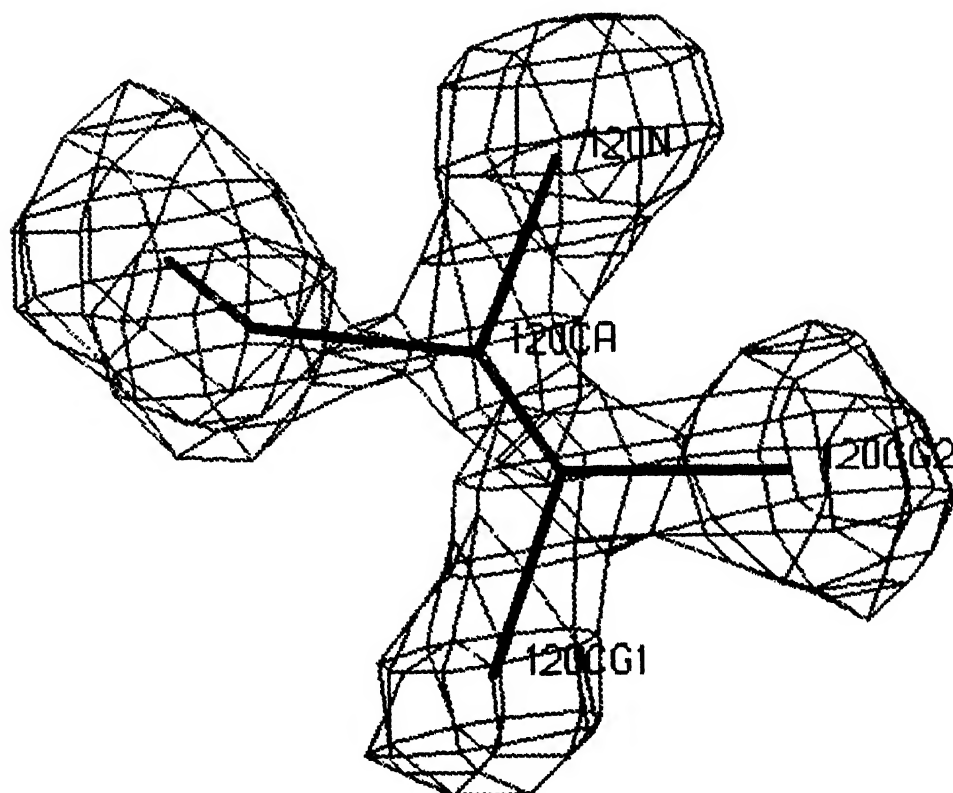


Fig 5 Isoleucine superposed in its electron density map as seen in a computer graph  
(Vinay Kumar and K K Kannan— *unpublished*)



crystallography, the atomic positions were determined from electron density maps stacked in three dimensions and protein models were then built. This was a laborious process and prone to errors. Professor F.M. Richards introduced the optical comparator method for easy electron density interpretation and model building. In this device illuminated three dimensional map sections are stacked behind a half silvered mirror and the atomic model known as Kendrew-Watson skeletal models were built by comparing the superposition of the reflected image of the model on to the map. For the past 15 years this has also been relegated to history as computer controlled real time graphics machines have steadily replaced these optical comparators. The computer graphics has made electron density interpretation and model building at once a pleasure at the same time providing one with very accurate coordinates at one go.

Indeed the whole of protein crystallography is computer intensive. Today for a successful practice of protein crystallography, a laboratory requires to be equipped with upto date X-ray crystallography equipment, good biochemical and molecular biology facility, access to high speed computers, real time high speed computer graphics facilities and state of the art computer software. Protein crystallography groups in the world are among the few to use super computers or parallel processors for their computational requirements. Our group was one of the very first users of the 8 node high speed parallel processor with 80 Megaflops rating, developed indigenously at BARC.

The structure of the protein is the base for understanding its function and inhibitor/drug interactions. Once the structure is determined such studies can be planned and undertaken and the results interpreted in a meaningful way. The rest of the lecture illustrates this with the work I and my colleagues have carried out at BARC on carbonic anhydrase isoenzymes.

## STRUCTURE AND FUNCTION OF CARBONIC ANHYDRASE

Carbonic anhydrase (E.C. 4.2.1.1 carbonate dehydratase) is one of the fastest enzymes catalysing the interconversion of  $\text{CO}_2$  and water to bicarbonate and a proton. The eventual proton transfer in this reaction at the rate of a million per second per enzyme molecule has engaged the skills of many scientists. In 1961, human carbonic anhydrase II (HCA II) was purified by ion exchange chromatography and electrophoresis and

crystallized by Bror Strandberg, Anders Liljas and his colleagues in Uppsala Sweden. Crystals of purified HCAI were obtained by me and my colleagues in Uppsala in 1967. The biochemistry and kinetic studies were carried out by Sven Lindskog and Per Olof Nyman at Gothenberg, Sweden and by Edsall and his colleagues at Harvard university, USA. The early years were one of intense collaboration between the groups with yearly meetings to share new results and also to discuss difficulties encountered in the research. This was a very stimulating period for a physicist like me to learn some of the vocabulary of biochemistry and enzymology and share the excitement of research with international forum of scientists.

Very few protein crystallography groups existed in the world at that time and the Swedish group was one of the pioneers in the field. Today protein crystallography is practised not only in educational institutions and research laboratories but very actively in many biotechnology and drug companies. Alas some of the research in this field is getting to be even secretive due to the involvement of these companies.

X-ray diffraction data were collected with low power x-ray generators on films using precession cameras. The method was very time consuming and required large number of crystals, of the order of 100 crystals per data set. The films were digitized by semi automatic or manual microdensitometers in the beginning. There were many thousand measurements to be made for every film. The spots conventionally called a reflection, are to be measured for determining their varying grey levels. Approximately 150 such films were measured for the native and each of the heavy atom derivatives to get all the grey levels, known as intensities of the reflections, and merged together for each compound with computer programs. It is indeed remarkable that all these were done in Uppsala by us with first generation computers FACIT and BESK which were 4K memory thermionic valve operated ultra slow computers.

In BARC, we have set up a protein crystallography laboratory with a high power rotating anode x-ray generator, precession camera, Arnodt-Wonacott oscillation camera for collecting intensity data from protein crystals, an automatic high speed microdensitometer controlled by a PDP11/34 computer to measure the films. While the FACIT and BESK computers occupied enormous amount of space, more than one floor, the faster PDP11/34 with more memory (256K) occupies only one cabinet space. Even the very modern parallel processing super computer developed

in BARC with atleast at 3 orders of magnitude higher speed, compared even to the PDP11/34 takes the same floor space as the PDP11/34. Even the modern day sleek PC486 is about 2 orders of magnitude faster than the PDP11/34. New developments like multiwire 2 dimensional detectors and image plate systems have replaced the film and the microdensitometer which have improved data collection efficiencies to a matter of a few days compared to months with the film method. The Molecular Biophysics Unit at the Indian Institute of Science, Bangalore has a multiwire area detector as a national facility for protein crystallographers. Storage ring synchrotron radiation x-ray sources has added many orders of magnitude power for protein crystallography research. Such a facility is available in a few countries, though an Indian synchrotron facility being built by the Centre for Advanced Technology, Department of Atomic Energy, Indore is expected to be available only around 1996. The synchrotron x-ray source has brought the possibility of real time enzyme kinetics from a realm of dream to reality. Laue technique of data collection which marked the advent of x-ray crystallography, has already been used successfully in a few studies of enzyme catalysis in protein crystals. We have used all these techniques except the Laue method for data collection for carbonic anhydrase. We have also established all the soft ware required for protein crystallography at BARC and successfully used them.

A reasonably good biochemical facility established recently has enabled us to purify and crystallize proteins such as HCAI and BCAII on a regular basis. This integrated approach we have adopted in BARC has given us the flexibility to take up protein crystallographic projects of our choice either on our own or in collaboration with other divisions or institutions.

HCAI was crystallized from 2.3M Ammonium sulphate, .05 Tris HCl at pH 8.5 by the novel method of seeding in glass capillaries. Four heavy atom derivatives were found suitable for the work after extensive screening of a large number of different heavy atom modifications and used in the MIR method of structure determination. The structure was built in an optical comparator, affectionately called "Richards Folly", and the atomic coordinates measured manually. After a large number of computations and computer graphics application the final structure was obtained which is very good by international standards.

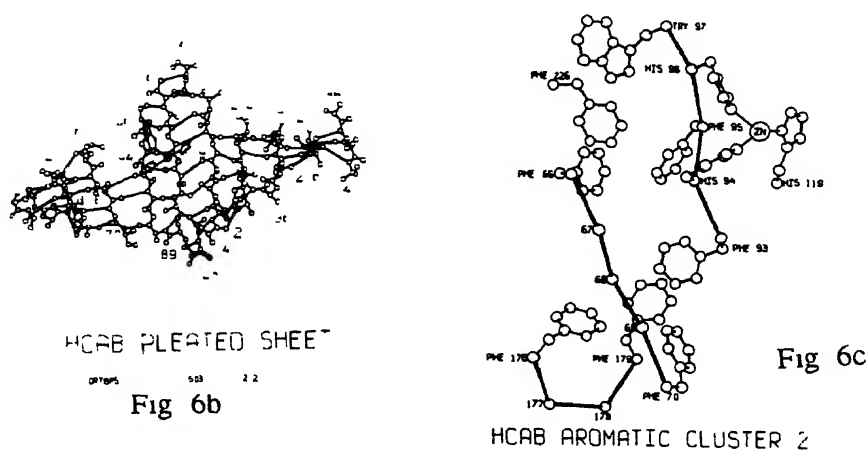
HCAI is a globular molecule, 41x42x55 Å in dimension, and built up of 10 central twisted pleated sheets and two minor plated sheet segments.



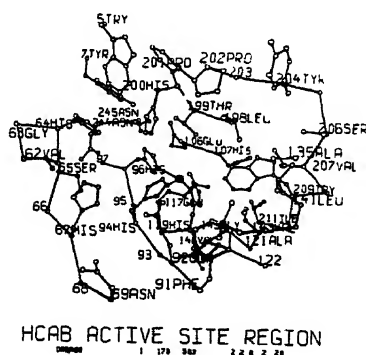
Fig 6a Ribbon Drawing of Carbonic anhydrase showing Zinc and the 4th coordinated water.

12-16% of the aminoacids are in 6 short  $\alpha$  or  $3_{10}$  helices. Four segments of the central plated sheet and two loops shape the active site cavity with an essential Zn ion bound firmly to three histidyl residues, His 94, His 96 and His 199 from the central pleated sheet segments. The cavity is funnel shaped with the zinc ion at its apex at 12 Å depth. A number of important

amino acid residues Tyr7, His64, His67, Asn 69, Glu92, Glu106, His200 provide a hydrophilic environment, while Phe91, Ala121, Leu131, Ala135,



- Fig 6b Extensive pleated sheet in HCAI an important secondary structure that stabilizes the structure and also provides the specific binding site for the essential zinc
- Fig 6c Aromatic/hydrophobic packing in HCAI which is invariant in carbonic anhydrases that is important for providing the globular character of the protein and also stabilizes the structure



molecules. It is interesting that Thr199, Glu106, Tyr7 are all conserved including the water molecules participating H bonding that are located in their vicinity. This definitely agrees well with the mechanistic proposal I had made several years back, suggesting for the first time the involvement of Thr199 and Glu106 along with the Zinc bound water molecule in the mechanism of action of carbonic anhydrases. The extension of the H-bond network to His 64, 67 and 200 is in line with the importance of these residues that recent site directed mutagenesis experiments have revealed.

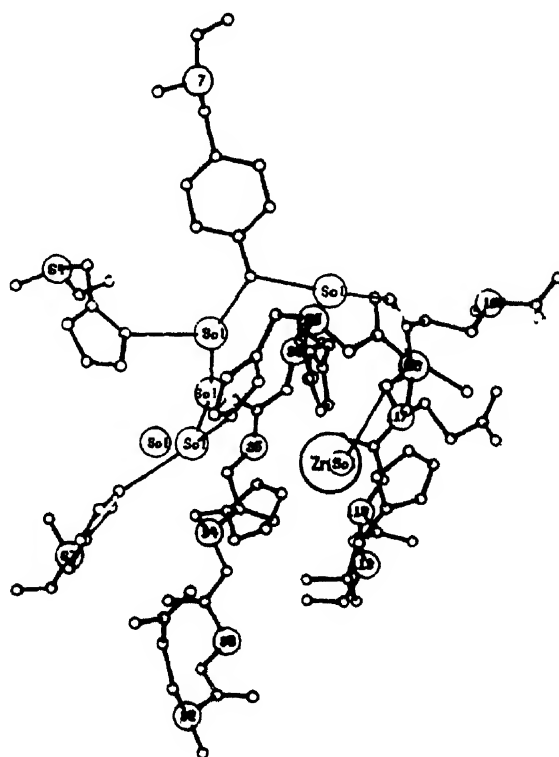


Fig 8 Active site of HCAI showing the extensive hydrogen-bonding where a number of invariant water molecules are also involved.

Inhibition studies of enzymes reveal a great deal about the function of these important biomolecules. Carbonic anhydrases are inhibited by a number of monovalent anions such as  $\Gamma^-$ ,  $\text{Br}^-$ ,  $\text{SCN}^-$ ,  $\text{CN}^-$ , etc. and also by the product of catalysis  $\text{HCO}_3^-$ . Aromatic and heterocyclic sulphonamides with unsubstituted sulphomido group are very specific and powerful inhibitors of CA's. In fact acetazolamide under the pharmaceutical name Diamox is used routinely as a medicine for treating Glaucoma, an eye

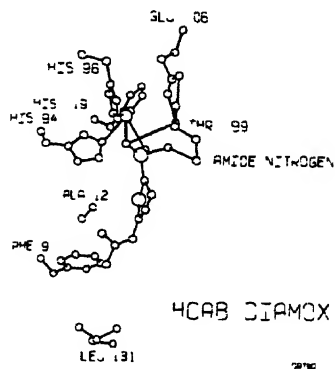


Fig 9

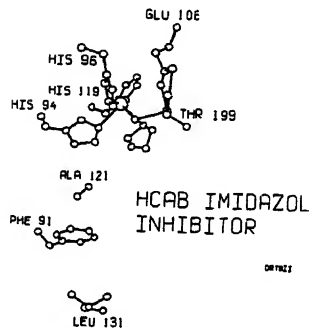


Fig 10

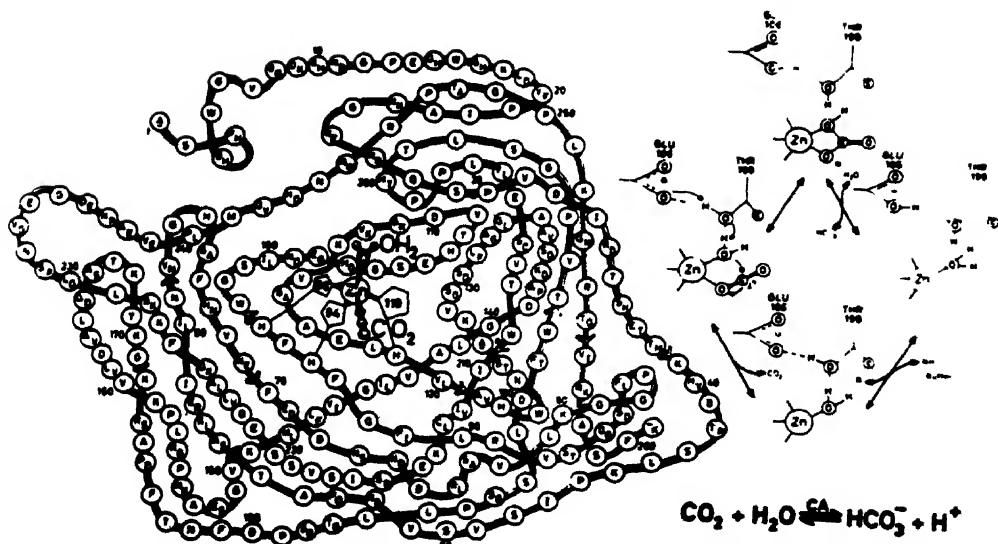


Fig 11 A schematic of the structure of carbonic anhydrases along with the mechanistic proposal wherein carbon dioxide is bound as a weak 5th ligand to zinc.

disorder. We have investigated the complexes of HCAI- $\Gamma^-$ , HCAI- $\text{HCO}_3^-$ , HCAI- $\text{AuCN}_2^-$ , HCAI-acetazolamide, HCAI-methazolamide as well as HCAI-benzene sulphonamide and determined their three dimensional structures to a high degree of precision. It is clear from these studies, that except for  $\text{AuCN}_2^-$ , all the other inhibitors bind to Zn by replacing the fourth coordinated water molecule, thus blocking the site of the activity linked group which also is the second substrate in the hydration reaction. By binding to the fourth coordination site these inhibitors compete with the binding site of the substrate  $\text{HCO}_3^-$  in the reverse dehydration reaction. Thus the inhibition of the CA's is explained very nicely from the crystallographic studies of HCAI- inhibitor complexes.  $\text{AuCN}_2^-$  binds indirectly to  $\text{Zn}^{2+}$  without replacing the 4th coordinated water molecule. In fact the mode of inhibition of this uncompetitive inhibitor, seems to be by disrupting the H bond between  $\text{Zn-OH}^-$  and Thr 199 thus delinking the important proton transfer pathway in the enzyme. The CN is also found to Hydrogen bond to the  $\text{Zn-OH}$ , which in fact is the only observed link in the structure of this inhibitor complex which seems to explain the remarkably good binding constant to the enzyme. This binding site also overlaps with a 5th coordination site that was observed in the structure of HCAI-imidazole complex. Imidazole is the only competitive inhibitor of  $\text{CO}_2$  hydration reaction known for HCAI, all the others are competitive inhibitors of the dehydration reaction. Structures of HCAII complexed with a number of anion inhibitors have been studied by Liljas and his colleagues in Lund, Sweden, These have also revealed that Zn in carbonic a hydase has potential for flexible liganding directions and can accommodate 4 or 5 coordination. It is indeed gratifying that my mechanistic proposal which involved weak binding of  $\text{CO}_2$  as a 5th distant ligand to  $\text{Zn}^{2+}$  with the fourth coordinated OH oriented and positioned favourably to attack CO to produce the product HCO by enzyme catalysis has got good support from these inhibition studies.

The most important support for my mechanistic proposal has also come from site directed mutagenesis studies of Thr199 and Glu106 in HCAII in Sweden and USA. The substitution of Thr199 with other amino acids abolishes the activity of the enzyme. Substitution of Glu106 with other than Asp106 also abolishes the activity and thus bringing out the importance of these two aminoacids which i had invoked in the mechanism of the enzyme. Mutagenesis of His64 in HCAII does not abolish the activity, especially when imidazole is used as buffer. His64 is thus another important residue in HCAII. Our recent structural studies, similarly shows



that His67 and His200 are important for HCAI. Site directed mutagenesis work underway in BARC will throw considerable light on the importance of these residues.

## CONCLUSION

Structural studies of HCAI in BARC and the comparison of the results with those from other laboratories on HCAII reveal a remarkable similarity in the three dimensional structure of carbonic anhydrases. In fact our work on BCAII and the work elsewhere on BCAIII by Liljas in Uppsala, Sweden, also shows that this similarity is universal for all the mammalian carbonic anhydrase isoenzymes. Structure of inhibitor complexes of CA's in relation to the biochemical studies have provided insight into the mode of inhibition. The drug interaction studies of sulphonamides with CA's has also given a new tool for rational drug design in particular for Carbonic anhydrases. A lot more can be talked about these fascinating enzymes, though I have given only a glimpse of the possibilities of protein crystallography in this short lecture.

## ACKNOWLEDGEMENT

It is a pleasure to thank Bror Strandberg who introduced me to protein crystallography, encouraged me to take up the challenges in the field and brought me up in the initial years of this fascinating field of protein crystallography, Anders Liljas, with whom I have had many years of positive collaboration, Kerstin Fridborg and Seved Loughren who taught me the purification of CA's during my stay at Uppsala. To Dr. Chidambaram, who provided me the opportunity to build and nurture the protein crystallography group at BARC I owe a very special thanks. The contribution of my colleagues, Ramanadham, Vinay Kumar, Chakravarty, Yadava, Padma Satyamurthy to the groups activity on carbonic anhydrase in BARC which made his talk possible. My sincere thanks are due to them. I want to thank INSA for awarding the Jawaharlal Nehru Birth Centenary lectureship in Biological sciences and the Lucknow local chapter of INSA for the hospitality shown to me and also for holding this lecture at Lucknow.

## Bibliography

1. Fridborg, K., Kannan, K. K., Liljas, A., Lundin, J., Strandberg, B., Strandberg, R., Tilander, B. and Wiren, G. *J Mol.Biol*, **25**, (1967) 505-516.
2. Lindskog, S., Henderson, L.E., Kannan, K.K., Liljas, A., Nyman, P.O. and Strandberg, B., in *The Enzymes* 3rd edition, (Ed. Boyer, P.D.) Academic Press, New York, **5**, (1971) 587-665.
3. Kannan, K.K., Notstrand, B., Fridborg, K., Lovgren, S., Ohlsson, A. and Petef, M. *Proc Natl Acad.Sci U S A*, **72**, (1975) 51-55.
4. Liljas, A , Kannan, K.K., Bergsten, P.C., Waara, I., Strandberg, B , Carlbohm, U., Jarup, L., Lovgren, S and Petef, M., *Nature New Biol* **235**, (1972) 131-137.
5. Kannan, K.K , Liljas, A , Waara, I., Bergsten, P.C., Fridborg,K., Lovgren, S., Strandberg, B.,Bengtson, U , Carlbohm, U , Jarup,L., and Petef, M., *Coldspring Harbor Symposia on Quant. Biol* , **36**, (1971) 221-231.
6. Kannan, K K., Waara, I., Notstrand, B., Lovgren, S., Borell, A., Fridborg, K. and Petef, M., in *Drug Action at the Molecular Level* ed. Roberts, C G K., Macmillan Press, U K, (1977) 73-93.
7. Chakravarty, S., Yadava,V. S , Vinay Kumar, and Kannan, K.K , *Biosci* **8**, (1985) J491-498.
8. Vinay Kumar, Kannan, K. K. and Padma Sathyamurthy, *Acta Cryst D* , (1993) (in press)
9. Vinay Kumar Kannan, K. K. and Chidambaram, R., *Curr Sci*. **58**, (1989) 344-348.
10. Kannan, K K., In *Molecular Conformation and Biological Interactions* (ed.) Balaram, P, and Ramaseshan, S, Indian Academy of Sciences, Bangalore, (1991) 131-147.
11. Kannan, K K., Petef, M., Fridborg, K., Hilda Cid-Dresdner and Lovgren, S. *FEBS lett* , **73**, (1977) 115-119.
12. Kannan, K K., in *Molecular Biology, Biochemistry and Biophysics*.Advanced Methods in Protein Sequencing., (Ed Needleman, S B, Springer Verlag, Germany, **25** (1977) 77-185
13. Kannan, K. K and Chidambaram, R, *Proc Ind Acad Sci* , (Chem. Sci.), **92**, (1983) 497-507
14. Blundell, T L. and Johnsson, L. N., *Protein Crystallography*, Academic Press, (1976) New York.
15. Jarup, L., Kannan, K K., Liljas, A. and Strandberg, B., *Computer Programs in Bio Med* , (1970) 1 74-77.
16. Jones, T A., *J. Appl. Cryst.* **11**, (1978) 268-271.
17. Kannan, K. K., Ramanadham, M. and Jones, T A., *Ann. New York Acad. Sci.*, **429**, (1984) 49-60.



Viswanathan, Sasisekharan (b. 28 June 1933) did Ph.D. (1959) from University of Madras. He is Professor, Molecular Biophysics Unit, Indian Institute of Science, Bangalore; was formerly Professor of Physics, University of Madras (1971- 72); Visiting Professor, Frick Chemical Laboratory, Princeton University, USA (1970-71).

Sasisekharan has done significant work on the conformation of biopolymers, specially polynucleotides and polypeptides. Developed incisive methods for arriving at optimal conformation of macromolecules. On the basis of the methods he

has suggested a new alternative model for the Watson-Crick double helical structure of DNA.

Sasisekharan is Fellow of Indian Academy of Sciences. He is the recipient of Shanti Swarup Bhatnagar Prize (1978); FICCI Award (1981), J.C Bose Medal (INSA) (1983), ASTRA Chair in Biological Sciences (1985); Watumull Honor Summus Medal (USA) (1987); Fogarty Scholar in Residence, NIH (USA) (1988-90), Om Prakash Bhasin Foundation Award (1989); J.C Bose Award (Hari Om Ashram-UGC) (1985).

*Viswanathan Sasisekharan was elected to the fellowship of the Academy in 1980*

**Address:** Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, Tel. (080) 344411 (Ext) 458; Res: E- 12, New Housing Colony, IISc, Bangalore 560 012; Tel. (080) 349148

## **HISTORY OF LEFT-HANDED DNA**

V SASISEKHARAN

### **PREAMBLE**

A correct description of the structure of DNA is important not only for its own sake, but also because that would make clear the mechanism by which DNA can replicate and transmit genetic information. The well-known structure for DNA as initially proposed by Watson and Crick in the early fifties and subsequently refined by Wilkins and co-workers, is a right-handed double helix. Watson, Crick and Wilkins were awarded the Nobel Prizes for their work on the DNA structure. The original structure proposed by Watson and Crick was not quite correct in the sense that the structure was in poor agreement with the X-ray data. For a better fit with the X-ray data, Wilkins and co-workers slightly modified the original structure proposed by Watson and Crick but at the same time retained the essential features of the double helix. The original proposal of the double helix was based on the then available chemical and X-ray data. The chemical studies on DNA showed that it consisted of two long polymer chains. Each polymer chain is basically made up of several monomeric units, the nucleotides. The variation in the chemical structure of the nucleotide comes from the nature of bases attached to the sugar through the glycosidic bond (Fig. 1). The determination of the structure of DNA thus reduces to: (1) topology or conformation of each polynucleotide chain, and (2) spatial arrangement of the two chains: A proposed structure should be consistent with both chemical and X-ray data. Chargaff from extensive studies on DNA demonstrated that  $A/T \simeq 1 \simeq G/C$ . The gross feature of the X-ray fibre diffraction data is a cross pattern.

Considering first the topology of the chain the cross pattern in the fibre data was interpreted due to a helical structure of the polynucleotide. The question then arises how to spatially arrange the two chains about the helical axis. Fig. 2 is a schematic representation of the polynucleotide chain. The symbol B represents the bases, Adenine, Thymine, Guanine and Cytosine. S represents deoxy sugar and P the phosphate group. The helix axis can be located either to the left or to the right of the polynucleotide chain. Pauling chose the axis on the right of the chain (right upper half of Fig. 2) and this would result in a structure with

phosphates near the helix axis and bases away on the outside. Watson and Crick placed the helix axis on the left side of the chain as shown (Fig. 2). In this case, the phosphates will be outside and the bases inside. This scheme of arrangement alone can explain Chargaff data. For  $A/T \approx G/C$  can be readily achieved if specific base interaction is sought between A and T and G and C. The geometry of the bases are such that hydrogen bonds can be readily formed between A and T base pairs and G and C base pairs in a helical arrangement. It is readily seen that this structure suggests a possible copying mechanism for the genetic material and could explain the semi-conservative replication. For obvious reasons, the structure proposed by Pauling was incorrect. The chemical basis of replication, namely the specific base-pairing scheme of Watson and Crick has been well substantiated later by several studies including single crystal analysis of oligonucleotides.

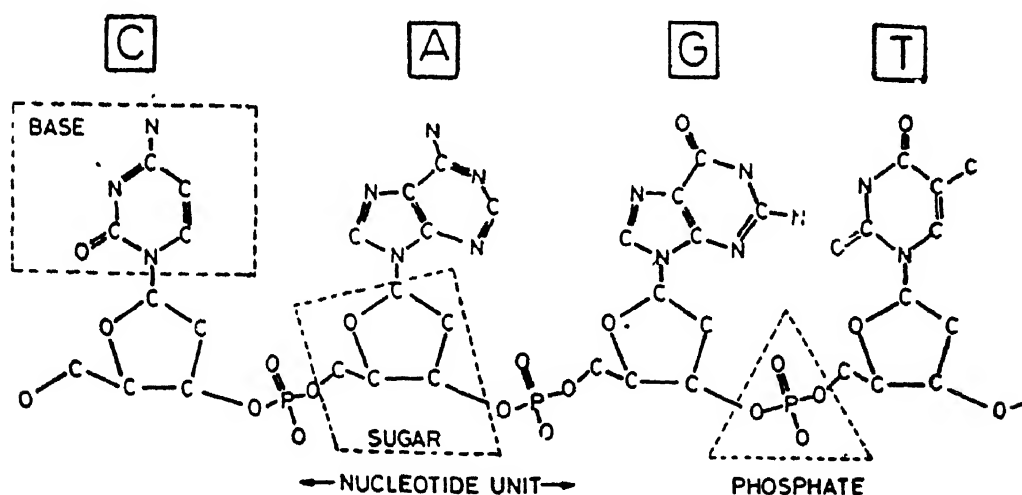


FIG 1 A polynucleotide chain (hydrogen atoms are not shown). The symbols C, A, G, and T stand for cytosine, Adenine, Guanine and Thymine.

The question then concerns with the topology of the polynucleotide chain backbone. As stated already, it was assumed that the cross pattern would be given only by a helical structure and not by any other structure. In such a case, one would like to know whether it is a right-handed one or a left-handed one: for right-handed and left-handed structures are not enantiomers. This is so because, the sugar unit in the nucleotide is a  $\beta$ -D Furanose (Fig. 3a). The enantiomer of this will be a  $\beta$ -L sugar (Fig. 3b). As the sugar in the polynucleotide chain is asymmetric, the right and left handed structures generated by linking successive sugar phosphate

backbones are not mirror images and therefore they lead to two different structures.

### THE DOUBLE HELIX

In the structure proposed by Watson and Crick (1953), each chain is a right-handed helix and runs anti-parallel to each other. The two chains as stated above are complementary to each other through the specific base-pairing schemes. Wilkins and co-workers studied the fibre diffraction pattern of DNA in great detail and showed that natural DNA fibres are usually dimorphic. They also showed that the dimorphic forms A and B are interconvertible in the solid phase, depending upon the conditions of the experiments namely, salt concentration and relative humidity. B-DNA

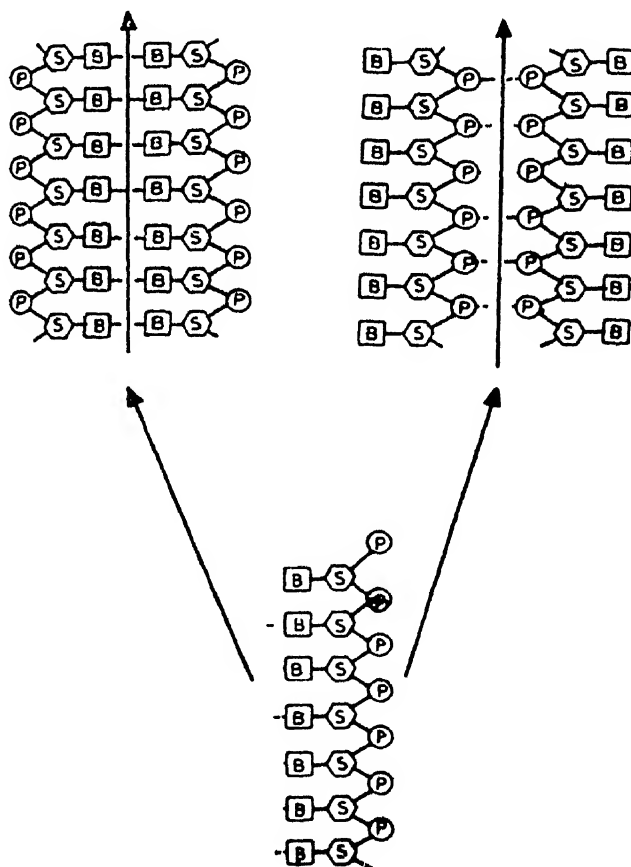


FIG 2 Schematic representation of the polynucleotide chain. The symbol B stands for the base, S for deoxy-pento-sugar and P for the phosphate. Location of the helix axis in the two different structures are also shown

is the form thought to closely resemble the living material and the structure proposed by Watson and Crick is for this form. A and B forms of DNA were interpreted in terms of right-handed double helices. Wilkins and co-workers were fully aware that from the poorly resolved fibre diffraction data, it was not possible to demonstrate clearly the handedness of the DNA i.e. whether DNA is right or left-handed. They observed that left-handed B-DNA was also possible; but found left-handed DNA in the A form was stereochemically unviable. It was then argued that both the forms are interconvertible at solid phase without any change of handedness. Hence, both A and B forms were taken to be right-handed (Fuller et al. 1965). Subsequently, studies on the different forms of DNA therefore, retained the right-handed sense of the double helix. Thus, for nearly two decades or so, the structure of DNA was taken to be a right-handed double helix.

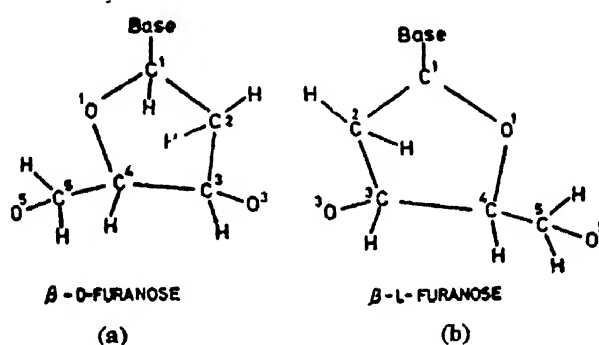


FIG 3a & b Sugar unit in the nucleotide of DNA is the  $\beta$ -D-furanose (a). The enantiomer of this is a  $\beta$ -L-furanose (b) not naturally occurring in DNA. Right- and Left-handed structures generated using  $\beta$ -D-furanose (a) are not mirror images and therefore lead to two different structures

### POLY (DLDC): A SUGGESTED LEFT HELICAL STRUCTURE

In 1970, Langridge and co-workers (Mitsui *et al.* 1970) suggested that the synthetic duplex DNA poly (dI-dC) had a left-handed double helical structure. They however, could not convert poly(dI-dC) designated as D-DNA (Not usually observed in natural DNA) into either a A-DNA or B-DNA. In addition, the circular dichroism (CD) spectra of the poly(dI-dC) was more or less exactly opposite to that obtained from B-DNA (Fig. 4). Therefore, they suggested that poly(dI-dC) was left-handed. It was about that time, I spent my sabbatical at Princeton, as a Visiting Professor in Bio-chemical Sciences. At Princeton in 1971, detailed fibre diffraction studies were carried out by me on poly(dI-dC). It was possible for me to

obtain a B-DNA form for poly(dI-dC) and to interconvert B and D forms of poly(dI-dC) by changing the relative humidity of the specimen. Feeling as did Wilkins group that such alterations were unlikely to have so drastic an effect as a change in the helical sense in the solid state, it was concluded by me that the helical sense of D-DNA, A-DNA and B-DNA must be the same. Given the argument of Wilkins group against the possibility of constructing a left-handed double helix for A-DNA it seemed to me that D-DNA too must be right-handed. The anomalous CD spectra of the D-DNA however, remained unexplained. Just as any one else, I thought that to explain CD spectra alone, a non Watson and Crick proposal for the structure should not be considered seriously. I had a discussion with Langridge about this and about the earlier studies on DNA, as he was one among those in Wilkins group, responsible for the detailed X-ray and model building studies on DNA. The outcome of this discussion was to raise a methodological problem in my mind. I realised that earlier studies had not systematically explored all the possibilities. The approach in common with the earlier ones of Wilkins and co-workers and with everyone else in this field seemed to have been that 'a model was built, Fourier transforms were calculated and if the transforms agreed with the diffraction data, the structure was taken to be correct.' The fit between the transforms and the X-ray diffraction evidence was not so close that there could not be another structure consistent with the data. Fibre diffraction data of resolution of the order  $3\text{\AA}$  defy any direct determination of the structure of DNA in contrast to single crystal analysis at high resolution (of the order of  $1\text{\AA}$ ) which allows determination of the molecular detail of the structure. In the light of these considerations, I resolved to begin a systematic investigation of the models of the conformation of DNA possible given the constraints of X-ray fibre diffraction evidence.

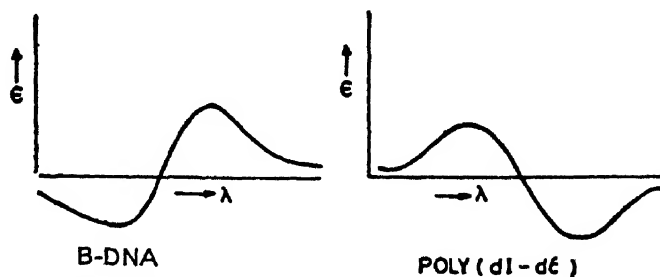


FIG 4 CD Spectra of B-DNA and poly (dI-dC). Note that the spectra are more or less opposite to each other.



## MODEL BUILDING STUDIES : A COMPARISON BETWEEN POLYPEPTIDES AND POLYNUCLEOTIDES

It may be of interest to point out here the main differences in the development of the model building studies of polypeptides and of polynucleotides. In the case of polypeptides, the model building studies were carried out by precisely obtaining the geometry of the monomer unit, namely, the peptide unit, from the analysis of single crystal data. In the case of polynucleotides, the geometry of the nucleotide unit was not fully understood and the model building studies proceeded simultaneously with the determination of crystal structures of nucleosides and nucleotides. Even when the refinement of the polynucleotide structures was made against X-ray fiber diffraction data by Arnott and co-workers (Arnott 1970), only information from crystal structures of nucleosides and nucleotides were available. However, such data provide only information about the stereochemistry of nucleotides and not the relative orientation of the nucleotides around the 3' and 5'- phosphodiester bonds. The relative orientation is the key factor which determines the secondary structures of nucleic acids and details of this orientation can be obtained only from single crystal studies of dinucleosides monophosphates and other higher oligomers. Even after almost two decades following the proposal of the double helix, only a few structures of oligomeric DNA have been solved at molecular resolution. The paucity of such structural data called for an altogether independent and complementary approach in the study of the plausible conformational variants of DNA. The approach was model

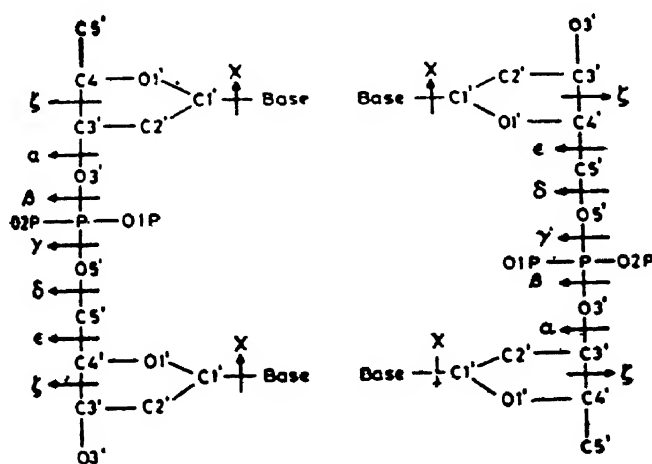


FIG 5 Torsion angles of the base paired dinucleoside monophosphate chosen as a repeating unit

building studies through stereochemical considerations. Initially, the flexibility of the furanose ring was studied in great detail while still in Princeton (Sasisekharan 1973). Subsequently on return to India in Bangalore a systematic analysis was made on the then available crystallographic data, mostly of dinucleoside monophosphates and a couple of higher oligomers and their features obtained (Sasisekharan & Pattabiraman 1978).

### THE FLEXIBILITY OF NUCLEOTIDE AND STEREOCHEMICAL GUIDELINES: RIGHT-AND LEFT-HANDED DUPLEXES

The data revealed certain correlations among the major torsional degrees of freedom present in the structure (*see figure 5 for nomenclature of torsions*). *For the first time it was recognised by us that an enormous degree of conformational flexibility is inherent in the nucleotide unit, the building block of the polynucleotide duplex, and at the junction of two neighbouring nucleotides in polymers.* This flexibility was made use of extensively in our building of molecular models, studies that were done following stereochemical guidelines. Thus, it was possible not only to identify the stereochemical nature of this flexibility but also to show that both right and left-handed structures of DNA, some of which were subsequently seen in single crystals, followed as a natural outcome of the flexibility (Gupta *et al.* 1980a). Thus, we discovered that there are no stereochemical grounds preventing the construction of a left-handed helix and that the rigidity imposed by earlier workers on the conformation flexibility of the nucleotide unit was artificial.

Initially, using the stereochemical guidelines, double helices with mononucleotides as the repeating units were investigated (Gupta *et al.* 1980a, Sasisekharan *et al.* 1981b). This led to the elucidation of left-handed DNA duplexes, hitherto ignored, and at the same time significantly improved upon the understanding of the existing right-handed models. These duplexes were compatible with the fiber diffraction data of various forms of DNA and did not permit discrimination between right and left-handed models (Fig. 6).

Subsequently, molecular conformations of DNA with alternating purine and pyrimidine sequences were studied (Sasisekharan *et al.* 1981a). It was observed that there could be a variety of sequence-specific and conformationally distinct right-handed and left-handed duplexes.

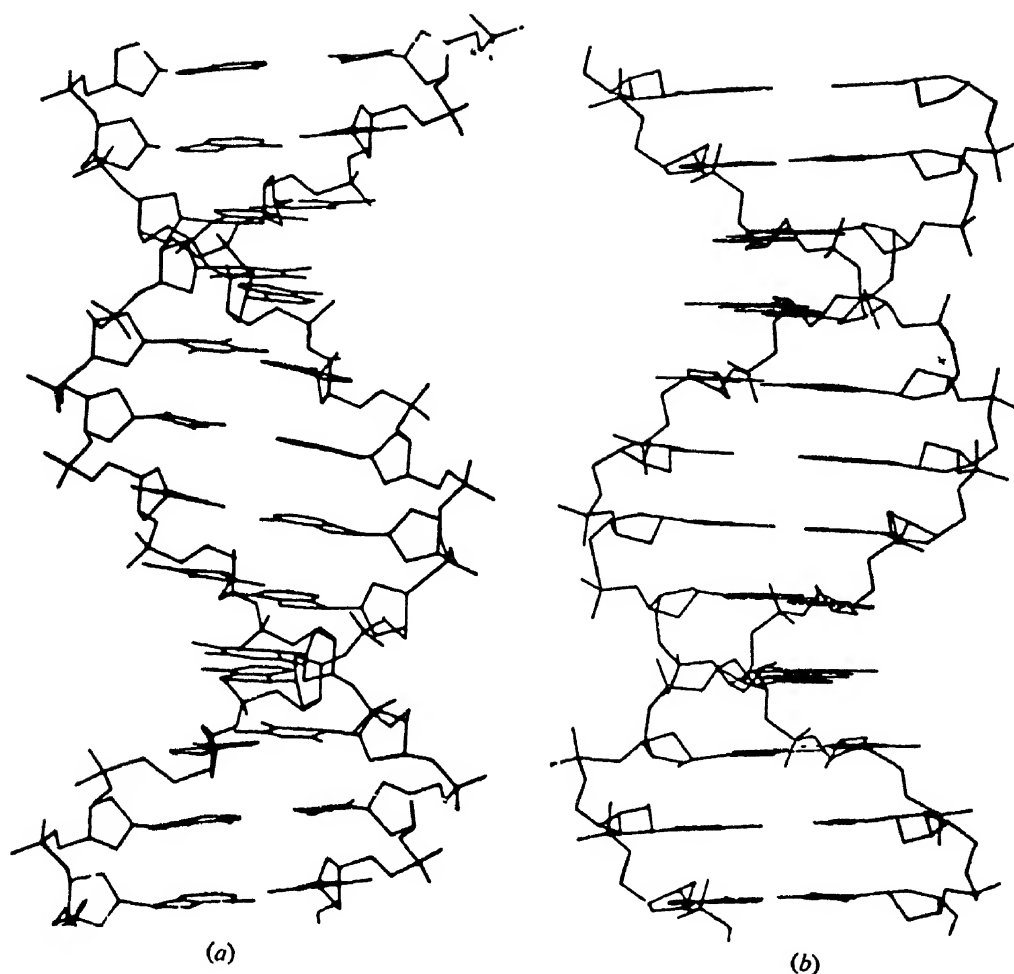


FIG 6 Models (Line drawings) of (a) right-handed and (b) left-handed duplex DNA agreeing equally well with the diffraction data of B-DNA

### THE RL MODEL FOR DNA

The possibility of DNA in either handedness raised the question whether one could join a segment of a right-handed structure with a segment of a left-handed structure. It was indeed shown to be possible due to the conformational flexibility of the sugar phosphate backbone (Sasisekharan & Pattabiraman 1976, Gupta *et al.* 1980c). The link possessed the allowed stereochemistry of the backbone and retained classic Watson-Crick base pairing. Bases however, had a novel kind of stacking arrangement called the inverted stacking which was shown to be energetically as favourable and also as frequently observed in single crystals as the normal type (Gupta & Sasisekharan 1978). The joining of the right-and left-handed

double-helical fragments resulted in a new structure of DNA called the RL model (Fig. 7). Thus, for the B-DNA, a structure with alternating right-helical and left helical segments of approximately 5 bp in a repeat of 10 bp was proposed (Sasisekharan & Pattabiraman 1976). Further details and implications of this structure and a comparison with the double helix are given elsewhere (Sasisekharan *et al.* 1977, 1978, Sasisekharan 1980, 1981).

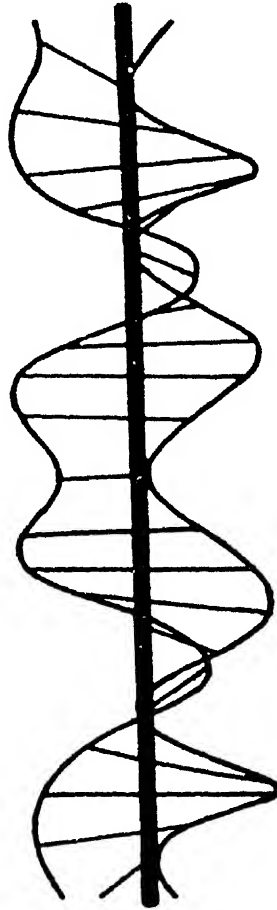


FIG 7 A schematic diagram of RL model. Segments of right-handed duplex (R) and left-handed duplex (L) could be uniform or zig-zag structures

### RL MODEL VS WATSON-CRICK DOUBLE HELIX

The essential differences between the double helix of Watson and Crick and a RL model are as follows: (i) In the double helix, each chain is a right-handed helix while in a RL model each chain consists of alternating right-and left-helical fragments. (ii) Taking B-DNA as an example, in the double helix the 1st and 11th base pairs are related by one full turn (i.e.

360°) while in a RL model, they are only related by a rotation of 0° to 360°. In an ideal RL model for B-DNA, the net coiling for ten base-pairs is 0°. But due to flexibility at the link, a finite coiling (twist up to 360°) is possible. For example, an ideal double helix would predict a net number of coiling (linking number  $L_k$ ) as 500 for a DNA of length 5,000 base pairs while for RL model  $L_k$  should be within 50. Thus, a RL model makes uncoiling during replication easy to visualise which had always been a bothersome aspect of the double helix. It may be mentioned that purely from the consideration of replication Rodey et al. (1976) put forward a model similar to the RL model proposed by us. However, their model had normal stacking of bases at the link. Such a possibility was also earlier considered by us but we preferred the present model because of its stereochemical superiority over the other. Our approach to the problem was distinctly different from that of Rodley et al. We first recognized the conformational flexibility inherent in DNA, formulated a stereochemical guideline to exploit the same and arrived at the RL model. It thus happened that replication was easy to visualize on the basis of the RL model. Rodley et al., on the contrary, considered uncoiling of the two strands during replication as a great disadvantage for the double helix and therefore, put forward an alternative SBS model for B-DNA (their version of the RL model). They also argued that the X-ray data are in better agreement with their model rather than the double helix.

All naturally occurring closed circular DNAs do not separate out into two single circles on heating thus indicating a non-zero  $L_k$ . But this does not favour one model over the other unless one makes a direct measurement of  $L_k$  to check if it is in agreement with the double helix or a RL model. In 1979, in order to answer this question, Crick et al. did electrophoretic measurements with a few circular DNA duplexes. They concluded that their results could be interpreted only in terms of the double helix of DNA. When these results were reported, left-handed DNA fragments at atomic resolution were not known. Absence of such structures was also taken as an evidence against the possibility of a RL model and that DNA is a right-handed double helix only.

### LEFT-HANDED Z-DNA

Not before too long, in 1979 itself, interestingly, the first ever solved single crystal structure of short length of DNA double helix turned out to be left-handed. Alex Rich and co-workers solved the crystal structure of a hexanucleotide (CGCGCG) having alternating G and C bases. Though the

base pairing scheme is of the Watson-Crick type, the phosphates of the polynucleotide backbone trace a zig-zag left-handed roughly helical path and hence the structure was termed as Z-DNA. Again the next crystal structure, a tetranucleotide with alternating purine pyrimidine sequences (CGCG), solved by Dickerson and co-workers turned out to be similar and left-handed. Both these discoveries were of interest to us as our earlier investigations had indicated that the exact topology of the polynucleotide chains would be a function of the preferred conformations of the nucleotides of the purine and pyrimidine bases and these reports confirmed this (Figs. 8 and 9). (Gupta *et al.* 1980c).

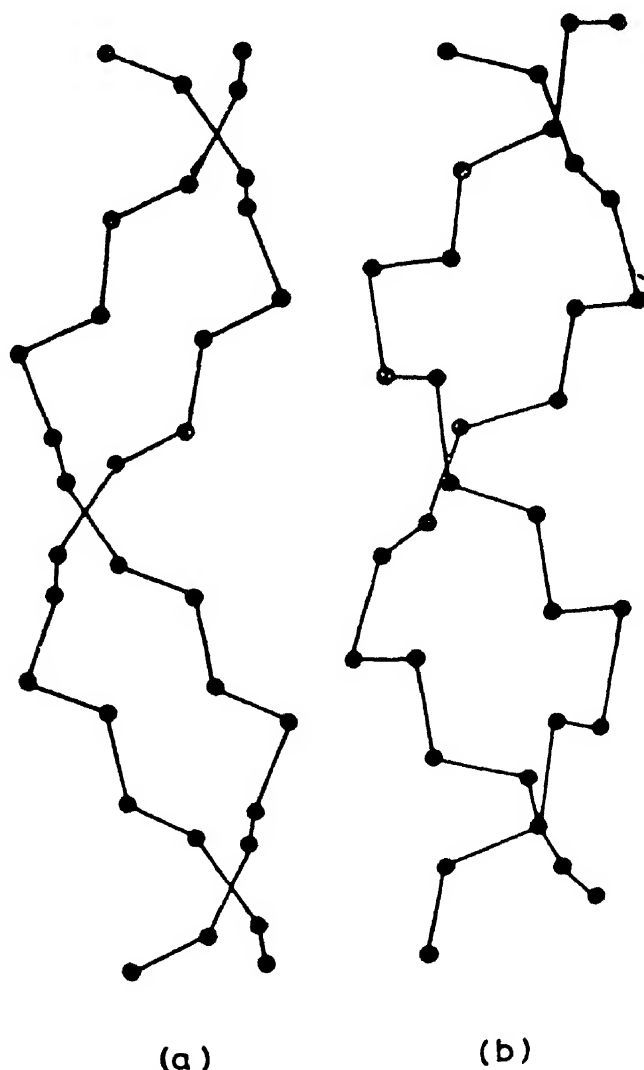


FIG 8 Comparison of left-handed zig-zag structures predicted by us with the gross crystal structures of  $d(CG)_3$  and  $d(CG)_2$  (a) left-handed zig-zag progression of the phosphate group (●) in the structure predicted (LZ1) by us; (b) Progression of the phosphate group in the  $d(CG)_3$  crystal (Z-DNA)

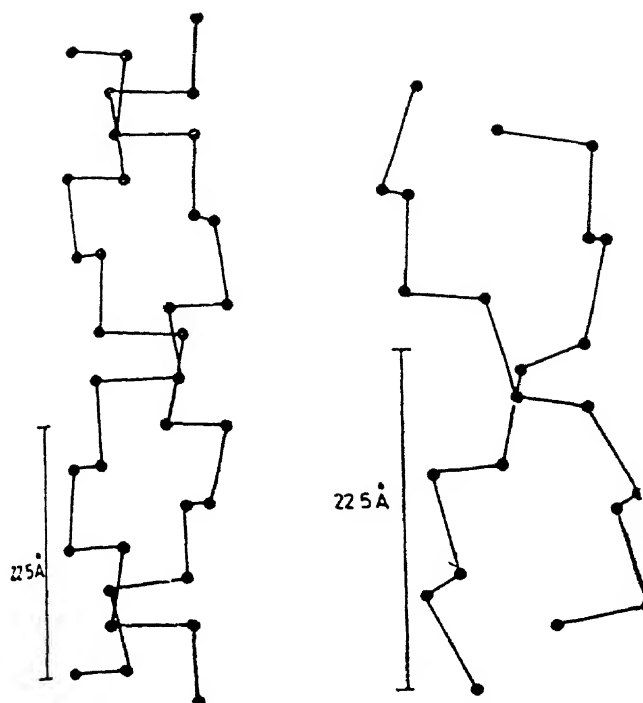


FIG 9 (a) Progression of the phosphate group (●) in the left-handed zig-zag (LZ2) helix predicted by us; (b) Progression of the phosphate group (Z'DNA) in the  $d(CG)_2$  crystal

Subsequently in 1980, Dickerson and co-workers (1980) observed a right-handed B-DNA like conformation in the crystal structure of the duplex of a dodecamer, CGCGAATTCGCG. It was observed that there could be conformational variants within B-DNA itself and the structure of the dodecamer is sequence specific.

Thus, in 1979-1980, the existence of both right-and left-handed structures in the fragments of duplex DNA was established in crystals. As early as 1972, the ability of the polynucleotide to assume a left-handed double stranded helical conformation in solution was suggested by Pohl and Jovin (1972) upon observing an inversion of the CD spectra of poly (dG-dC) under conditions of high ionic strength. Much later in 1981 it was confirmed by Raman spectroscopic studies that the poly (dG-dC) at high ionic strength and the hexanucleotide  $(CG)_3$  in the crystal have similar structure (Thamann *et al.* 1981). Thus, poly (dG-dC) is the only polynucleotide known whose structure has been interpreted in terms of a right-handed helical conformation under certain conditions of experiment and a left-handed helical conformation under certain conditions of experiment. During the last 4 years, extensive studies on the left-handed form of DNA have been carried out using NMR, Laser Raman, CD, ORD

and other physico-chemical techniques. Further it was also shown that  $(\text{dG-dm}^5\text{C})_n$  and  $(\text{dC-dA})_n$   $(\text{dG-dT})_n$  could also assume the Z conformation under appropriate conditions.

All of the nucleotide sequences mentioned above are present in genomic DNA. For example, alternating dG and dT residues are found in human, mouse, pigeon, xenopus, slime mould and yeast. The question remains whether such sequences when present in a heterogeneous sequence of DNA can exist in the left-handed Z-form in the cell. If so then a Z-DNA fragment has to co-exist by forming a stable link with a right-handed B-DNA thereby resulting in a RL segment in the DNA duplex, of the type proposed by us earlier. It would be of interest therefore to know how the left-handed Z-DNA or RL segment (junction of B and Z) could contribute to gene expression. Various experiments with antibodies specific for Z-DNA have given results consistent with the presence of Z-DNA in cells. Positive results have been indicated for polytene chromosomes for the macronuclei of *Stylonychia mytilus* and the nuclei of several types of rat cells. Although immunochemical studies yield reasonably reliable data there is the possibility that the anti Z antibodies are recognising macromolecules that share epitopes with Z-DNA rather than Z-DNA itself.

Several attempts are being made to isolate DNA fragment in the Z-DNA conformation and this would be difficult for in the process of isolation one might disrupt the Z-conformation. The studies are in progress to develop methods to stabilize the Z-form in DNA before isolation. During this period about 200 publications dealing with physicochemical, immunochemical and other studies on left-handed Z-DNA alone have been reported.

### REMARKS

Thus, it is clear that the structure of DNA is not entirely made up of right-handed double helix only as suggested by Watson and Crick. For the many functions that DNA has to perform it is obvious that duplex DNA should have a variable structure along its length. From what has been discussed above, it is clear that both right-and left-handed duplexes are plausible. It is therefore, very likely that segments of right-and left-handed duplex would be present giving rise to RL segments in DNA structure since a stable link can be made between them (Sasisekharan 1983). Two questions that arise then are what are the maximum lengths of continuous right-and



left-handed segments of DNA and under what conditions would these occur.

## References

1. Arnott S The geometry of nucleic acids; *Prog Biophys Mol Biol* 21 (1970) 265
2. Crick F H C, Wang J C and Bauer W R Is DNA Really a Double helix? *J Mol Biol* 129 (1979) 449
3. Drew H, Takano T, Tanaka S, Itakura K and Dickerson R E 1980 High salt d(CpGpCpG), a left-handed Z-DNA double helix; *Nature* 286 567
4. Fuller W, Wilkins M H F, Wilson H R and Hamilton L D Molecular configuration of deoxyribonucleic acid. IV X-ray diffraction study of the A-form; *J Mol Biol* 12 (1965) 60
5. Gupta G, Bansal M and Sasisekharan V Conformational flexibility of DNA: Polymorphism and handedness; *Proc Natl Acad Sci* 77 (1980a) 6486  
—, — and — Polymorphism and conformational flexibility of DNA: Right and left handed duplexes; *Int J Biol Macromol* 2 (1980b) 368  
—, — and — Reversal of handedness in DNA; A stable link between RU and LZ helices; *Biochem Biophys Res Commun* 97 (1980c) 1258  
— and Sasisekharan V Theoretical calculations of base-base interactions in nucleic acids. II Stacking interactions in polynucleotides; *Nucleic Acid Res* 4 (1978) 1655
6. Mitsui Y, Langridge R, Shortle B E, Cantor C R, Kodama M and Wells R D 1970 Physical and enzymatic studies on poly d(I-C), poly d(I-C), an unusual double-helical DNA; *Nature* 228 1166
7. Pohl F M and Jovin T M Salt-induced cooperative conformational change of a synthetic DNA: Equilibrium and Kinetic studies with poly (dG-dC); *J Mol Biol* 67 (1972) 375
8. Rodley G A, Scobie R S, Bates R H T and Lewitt R M 1976 A Possible conformation for double strand polynucleotides; *Proc Natl Acad Sci USA* 73 2959
9. Sasisekharan V 1980 A Possible alternative structure of DNA; in *Biomolecular Structure, Conformation, Function and Evolution*, Vol 1 p 352 ed R Srinivasan (Oxford: Pergamon Press)  
— Structure of DNA; *Curr Sci* 50 (1981) 107  
— Left-handed DNA duplexes; in the *Proc Cold Spring Harbor Symp. Quantitative Biol* Vol 47 (1983) p 45

- , Bansal M and Gupta G Structure of DNA re-examined; in *Structural Aspects of Recognition and Assembly in Biological Macromolecules* (1981a) p 501 ed M Balaban et al. (Jerusalem: Academic Press)
- , Gupta G and Bansal M Sequence-dependent molecular conformation of polynucleotides: Right and left-handed helices; *Int J Biol Macromol* 3 (1981b) 2
- and Pattabiraman N Structure of DNA predicted from stereochemistry of nucleoside derivatives; *Nature* 275 (1978) 159
- and — Double stranded polynucleotides: Two typical alternative conformations for nucleic acids; *Curr Sci* 45 (1976) 779
- , — and Gupta G An alternative structure of DNA and its relevance to DNA supercoiling; *Curr Sci* 46 (1977) 763
- , — and — Some implications of an alternative structure for DNA; *Proc natl Acad Sci* 75 (1978) 4092
10. Thamann T J, Lord R C, Wang A H J and Rich A The high salt form of poly (dG-dC) in left-handed Z-DNA Raman spectra of crystals and solution; *Nucleic Acids Res* 9 (1981) 5443
11. Watson J D and Crick F H C Genetic implications of the structure of Deoxyribonucleic acid; *Nature* 171 (1953) 964



**Naropantul Appaji Rao** (b. 25 September 1934) did Ph.D. (1959) from Rajasthan University. He is Professor at Department of Biochemistry, Indian Institute of Science, Bangalore

Appaji Rao's research deals with modulation of the activity of enzymes upon their interaction with regulatory molecules or among their subunits or with the external environment. These aspects are exemplified by his research on the regulation of nucleotide pyrophosphatase by AMP and aspartate transcarbamylase by UMP by hysteresis and by association-dissociation of the subunits caused by

binding of AMP. Also, he has made significant contributions to our understanding of the regulation of plant and fungal glutamine synthetase. His recent research has established the role of serine hydroxy-methyltransferase in normal and neoplastic tissues and this work led to the discovery of active site directed chemotherapeutic compounds effective against methotrexate resistant tumour cell lines. His research in the past 10 years has exploded the myth that continued inbreeding in human populations leads to deleterious genetic consequences.

Appaji Rao is Member (Secretary, 1965-69; Treasurer, 1970-75; Vice President, 1979-82, and president 1986-88) of Society of Biological Chemists, India. He is the recipient of Sreenivasaya Memorial Award (Society of Biological Chemists, India) (1980), J C Bose Medal Lecture (INSA) (1989).

**Address:** Professor, Department of Biochemistry, Indian Institute of Science, Bangalore 560 012.

## **GENETIC CONSEQUENCES OF INBREEDING IN A LARGE HUMAN POPULATION**

N APPAJIRAO

Inbreeding in human populations has fascinated sociologists and geneticists from a long time. Intuitively, we expect that recessive traits occur with increased frequency in the progeny of consanguineously related parents as they have a greater chance of inheriting a mutant gene from a common parent. Two individuals are said to be consanguineously related if they have atleast one common ancestor. The common ancestor must not be too remote or else the concept becomes meaningless. For convenience, in human genetics, ancestors more remote than great-great grand parents are not considered significant. Progeny of consanguineous parents are termed inbred. Genetic effects of inbreeding can be traced to the fact that an individual may carry a mutant gene inherited from a common ancestor. The modes of inheritance of these genetic disorders are classified as autosomal recessive, where a child inherits one mutant gene from each of the parent and is affected by the disorder; whereas the parents who have only one copy of the mutant gene are apparently normal. In autosomal dominant inheritance even if one copy of the gene is mutated the individual is affected by the disorder. In the X-linked recessive inheritance, as the females carry two copies of the X-chromosome, if one allele is mutated, they function as carriers and are apparently normal. As the males, have only one X-chromosome, if the gene is mutated he is affected by the disorder. In other words, females are carriers and transmit the mutant gene with 50% frequency to male offsprings who are affected by the disease.

One of the earliest studies on the effects of inbreeding was that of Darwin (who was himself an offspring of a consanguineous marriage), who studied the proportion of first cousins offspring among Oxford and Cambridge boating men and showed that there was 2.4% higher number of them among the non-consanguineous group (Darwin 1875). While endowed with great charm, these results of Darwin cannot be used to define the nature and extent of human inbreeding effects in man. Garrod (1902) showed that of the 18 cases of the inborn error of metabolism-Alkaptonuria, diagnosed in Europe and North America, 12 had been born

to parents who were first cousins. However, in order to maintain a sense of perspective on the level of morbidity, Garrod (1902) also reminded his readership of the extreme rarity of the condition among first cousins offspring as a whole, only 6 out of 50,000.

The practice of marrying close relatives is common among the populations of South India. Many text books, (Cavalli-Sforza & Bodmer 1971), on human genetics as well as Western social wisdom presume that consanguinity is confined mostly to areas of low population density and leads to decreased fertility, increased pre-reproductive loss, consequently to dwindling populations. Against this notion, in the 1981 census, the population of Karnataka, Andhra Pradesh, Tamil Nadu and Kerala, where consanguinity is extensively practiced, was 168 millions. It was suggested that continued inbreeding against a background of infectious diseases and malnutrition may have resulted in the elimination of deleterious genes from the population (Sanghvi 1966). My colleagues and myself therefore asked ourselves which of the two alternatives, (a) has inbreeding led to increased incidence of genetic disorders or (b) to cleansing of the gene pool, is correct?

Before I discuss these questions let me indicate the consequences of an autosomal recessive disorder. Due to an alteration in the gene either by mutation or deletion or other reasons, the product when it happens to be an enzyme could become inactive. For example, in Phenylketonuria, a disease caused by the deficiency of phenylalanine hydroxylase, the following consequences arise. Phenylalanine, which is derived from the diet is normally metabolised to tyrosine which is required for a variety of important products such as melanin, a pigment in the skin; thyroxine and epinephrine, which are hormones or neurotransmitters. When the enzyme is inactive, phenylalanine accumulates and the increased levels are channeled into toxic products. The combined effect of the deprivation of the metabolites arising from tyrosine and the toxic effects of degradative products of phenylalanine will lead to an irreversible brain damage and mental retardation.

One method to examine the validity of one of the questions raised in an earlier paragraph is to examine for the mutant gene among the mentally retarded individuals. Prior to our study, only a few of the disorders were identified in our population. It can be seen from the table 1 that several classical autosomal recessive disorders are present in the population as

indicated by their occurrence among the mentally retarded individuals (Sridhara & Rama Rao et al. 1977).

**Table 1**  
*Inborn errors of metabolism detected in a survey of 1280 mentally retarded children\**

| Disorder              | Number |
|-----------------------|--------|
| Phenylketonuria       | 14     |
| Maple syrup urine     | 5      |
| Hartnup's             | 4      |
| Homocystinuria        | 2      |
| Mucopolysaccharidosis | 3      |

\* Sridhara Rama Rao et al 1977

As one of our objectives was to examine the effects of inbreeding in humans, we examined the coefficient of inbreeding (F) which indicates the "Chance of an individual receiving a mutant gene from a common ancestor" and this value reflects the inbreeding level of the population. It was seen (Table 2) that the inbreeding coefficient in the mentally retarded individuals with a single gene defect (0.08) was higher than in the general population (0.07) (Rao & Inbaraj 1979). In the American population where marriage among close relatives is frowned upon, it is 0.001, (Freire-Maia 1968) a very low value compared to that in South India (Rao & Imbaraj 1979). One obvious and hasty conclusion that could be drawn from these results was that inbreeding leads to increased occurrence of mental retardation and increased rate of genetic disorders in the population. A moment's reflection would indicate some of the fallacies in this conclusion. One such fallacy arises from the fact that the samples were drawn from a highly preselected group with a bias in the referrals. It was therefore, necessary to substantiate this conclusion by examining the general population.

**Table 2**  
*Inbreeding coefficients among mentally retarded and general population*

| Disorder                                | Inbreeding coefficient |
|---|------------------------|
| Single gene: Mentally retarded          | 0.18                   |
| General population (Rao & Inbaraj 1979) | 0.02                   |
| American population                     | 0.001                  |

One interesting aspect of these genetic disorders of amino acid metabolism is in that the severity can be ameliorated to almost complete recovery in a few cases, by feeding the children a special diet lacking a particular amino acid. For this purpose it is necessary to begin the treatment before the symptoms appear and within a few weeks of birth. As presymptomatic screening for genetic disorders of amino acid metabolism has been demonstrated to be feasible in the developing countries and these disorders have been demonstrated to be present in our population (Sridhara Rama Rao et al. 1971) and most important they can be ameliorated, prompted us to presymptomatically screen the newborn and examine the relationship between inbreeding and the incidence of genetic disorders.

In addition to collecting three drops of blood for amino acid analysis, trained volunteers interview the mother and obtain information on consanguinity, religion, number of liveborn and living children, etc. This information was coded and stored on a computer at the Indian Institute of Science, Bangalore and analysed. The blood spot (0.6 cm diameter) was cut and eluted overnight into 50  $\mu$ l of ethanol and the extract was subjected to thin layer chromatography. The amino acids were located by reaction with ninhydrin. The separation of the amino acids were very good and 13-15 discrete amino acid bands were obtained (Ireland & Read 1972). The elevated levels of amino acid showed up strikingly as a thick band on the chromatogram indicating a possible genetic disorder. For example increased levels of phenylalanine is seen in phenylketonuria, tyrosine in tyrosinemia, leucine, isoleucine and valine in maple syrup urine disease, etc. Repeat analysis was done with the second spot. The child was followed up in more than 50% of the cases and confirmation of the disorder made biochemically (Appaji Rao N et al. 1988).

This programme was begun in 1980 and after initial teething trouble it has been going on steadily since 1982. The date (December 31, 1989) we have analysed more than 100,000 newborn children, drawn from 50 hospitals and Nursing Homes in Bangalore and Mysore (Appaji Rao et al. 1988).

Table 3 shows the type of pregnancies studied. There are a small number of twins and triplets and one quintuplet. Inbreeding contrary to expectation has had no effect on the twinning rate which is lower for the Indian population compared to Western cohorts (Radha Rama Devi 1981).

**Table 3**  
*Types of pregnancies surveyed in newborn screening programme*  
*(total 112, 339)*

| Type        | Number  |
|-------------|---------|
| Single      | 110,848 |
| Twins       | 747     |
| Triplets    | 11      |
| Quintuplets | 1       |

**Table 4**  
*Amino acidemias detected among 112, 369 newborns screened*

| Disorder                  | Number | One in |
|---------------------------|--------|--------|
| Phenylketonuria           | 6      | 18,728 |
| Tyrosinaemia              | 18     | 6,243  |
| Glycinaemia               | 7      | 26,053 |
| MSUD                      | 11     | 10,215 |
| Histidinaemia             | 3      | 37,456 |
| General amino acid amemia | 70     | 1,605  |

We detected 6 cases of hyperphenylalanenemia, 17 of tyrosenemia, 11 of branched chain aminoacidemia also known as maple syrup disease and 3 of histidinemia. The incidence of these disorders ranges one per 6,000 to 35,000 births. Tyrosinemia was found to be transient in many of these children and probably reflected prematurity of the baby as this enzyme was induced at later part of gestation (table 4). Generalised amino acidemia is not a genetic disorder and was caused by several unknown factors like infection, malnutrition, etc., and acts as a good internal control to compare inbreeding defects and the methodologies used. The incidence of single gene defects of amino acid metabolism in our population is very similar to that seen in several Western countries of Europe and America and in Japan. In all these countries it needs to be emphasised that there is very little inbreeding (Freire-Maia 1968). Table 5 shows the inbreeding coefficient of the children affected by single-gene disorders, generalised amino acidemia, in the newborn and the general population. The values are all over similar and it can be concluded that inbreeding as practiced in Karnataka has probably had no large effect on the incidence of these genetic disorders.



**Table 5**

*Comparison of the inbreeding coefficient (F) among the newborns with single gene defects, other newborns and general newborns*

| Group                                 | F      |
|---------------------------------------|--------|
| Single amino acid defects             | 0.0336 |
| General amino acidemia                | 0.0350 |
| General population (new born, Indian) | 0.0298 |
| American population                   | 0.0010 |

**Table 6**

*Consanguinity profile of the parents of the newborns screened for amino acidemias*

| Class              | Number | Percent |
|--------------------|--------|---------|
| Non-consanguineous | 66,609 | 62.3    |
| Uncle-Niece        | 18,240 | 17.1    |
| First Cousin       | 12,157 | 11.4    |
| Second Cousin      | 1,959  | 1.8     |
| Others             | 4,454  | 4.1     |
| Unknown            | 3,545  | 3.3     |

It has been generally postulated that urbanization and break up of the joint family system would have resulted in decreased consanguinity. However, the data in table 6 showed 17% uncle-niece marriages, 11% first cousin and 2% second cousin and about 8% other unclassified relatives. This large number of inbred offspring in our study provided a very fascinating material to study the effects of inbreeding on human populations. I have already indicated the effects of genetic disorders and I shall now turn to two other parameters.

Table 7 shows the number of liveborn and living children born to more than 100,000 mothers, the ratio between the live born and the living indicates the proportion of survivors. In a standard text book on human population genetics it is mentioned that inbreeding leads to loss of fertility and increased prereproductive losses (Cavalli-Sforza & Bodmer 1971). An examination of the proportion of survivors indicated that the inbreeding effect if any was marginal as the value for all the groups was nearly similar. A detailed and statistical analysis carried out indicated that there was no significant difference among the groups (Bittles et al. 1985).

**Table 7**  
*Number of live born and living children in the  
 non-consanguineous groups (Total: 111,593)*

| Class              | Liveborn | Living  | Proportion |
|--------------------|----------|---------|------------|
| Non-consanguineous | 153,797  | 147,893 | 0.961      |
| Uncle-Niece        | 44,417   | 42,537  | 0.957      |
| First Cousin       | 29,721   | 28,450  | 0.957      |
| Second Cousin      | 4,719    | 4,510   | 0.956      |
| Other, unknown     | 22,281   | 20,903  | 0.940      |
| Total              | 254,935  | 244,293 | 0.958      |

Table 8 shows the simplest of the statistical analyses, mean and the variance for the liveborn and the living. It is obvious from the figure that the mean number of children was slightly different among the groups. The mean number of children 2.2 shown in this table does not reflect the success of the family planning in Bangalore but indicates that the mothers are not at the end of the reproductive period. It was evident from these data that neither fertility (table 8) nor proportion of survivors (table 7) was seriously affected by inbreeding as practiced in this population (Bittles et al. 1985).

**Table 8**  
*Mean and Variance of live born and living children in relation to  
 consanguinity*

|                    | Live born |            | Living |            |
|--------------------|-----------|------------|--------|------------|
|                    | Mean      | Variance   | Mean   | Variance   |
| Non-consanguineous | 2.20      | $\pm 0.41$ | 2.12   | $\pm 0.38$ |
| Uncle-Niece        | 2.28      | $\pm 0.61$ | 2.18   | $\pm 0.53$ |
| First Cousin       | 2.35      | $\pm 0.66$ | 2.25   | $\pm 0.56$ |
| Second Cousin      | 2.32      | $\pm 0.59$ | 2.24   | $\pm 0.53$ |

The classical formula of Morton (Morton et al. 1956) which relates the survivorship and the proportion of deaths due to genetic and environmental causes has been extensively used to assess the effect

**Table 9**  
*Genetic disorders detected in 404 sick children*

| Type                  | Number | Percent |
|-----------------------|--------|---------|
| Single gene           |        |         |
| Autosomal recessive   | 11     | 2.7     |
| Autosomal dominant    | 24     | 5.9     |
| X linked              | 5      | 1.2     |
| Chromosomal           | 7      | 1.7     |
| Polygenetic uncertain | 16     | 5.6     |

**Table 10**  
*Inbreeding trends in parents of children with genetic disorders*

| Class            | Inbreeding (%) | Coefficient (F) |
|------------------|----------------|-----------------|
| Total group      | 48.7           | 0.04            |
| Single gene      | 45.2           | 0.05            |
| Autosomal        | 60.9           | 0.06            |
| General new born | 32.2           | 0.03            |

of inbreeding. It is obvious (figure. 1) that the points crowded close to the Y-axis were extrapolated assuming a linear relationship to indicate the effects of inbreeding. It is not necessary for me to highlight the hazards of such an extrapolation. Moreover, if we incorporate confidence limit of 5% into the data, the curve reaches ridiculous values when the inbreeding coefficient becomes large (Bittles & Makov 1985).

Bittles and Makov (1985) transformed this equation in different ways and analysed several models with different types of error distributions. An analysis of 20 different studies which have indicated strong inbreeding defects turn out to be mostly non-significant and many of the major ones become non-significant at 5% confidence limits in all the models.

One general conclusion that can be drawn from the study of the newborn is that considering the total population, inbreeding may have had no serious deleterious consequences. In human population studies or in any studies involving the human, the answers are not always black and white.

To illustrate this point I present the results of our study (Radha R D 1981) on sick children referred to us for diagnosing genetic disorders. Of the 404 cases examined, there were 35 single gene defects and we also saw a few chromosomal abnormalities and polygenic disorders (table 9). The predominance of single gene disorders probably indicates the bias of the referrals rather than any indication of true incidence. In view of our work on autosomal recessive disorders, there was a selection of patients suffering from these disorders. Small number of chromosomal abnormalities seen by us would indicate the preference of the clinicians to refer these patients to another centre in Bangalore. Table 10 shows the role of inbreeding in the etiology of these disorders. The inbreeding coefficient of this study group is slightly more than determined for the population as a whole. Although this observation could indicate the effect of inbreeding, it also might reflect the prejudice of the clinicians to suspect genetic disorders more often in consanguineous marriages than in other cases. This study confirms that many genetic disorders do exist in the population and that as environmental causes of childhood morbidity and mortality decline, the proportion of diseases with an exclusive or partial genetic etiology will increase. This phenomenon has been seen in many Western Societies.

The reemergence of malaria and the partial protection offered against this disease by glucose-6-phosphate dehydrogenase (G6PD) deficiency is offset by the drug-induced hemolytic anemia, these individuals suffer (Luzzatto 1985). G6PD is useful polymorphic genetic marker and the preferred pattern of uncle-niece and maternal cousin marriages could lead to increased incidence of homozygosity in the X-chromosome. For these reasons, it was of interest to examine the incidence of G6PD deficiency in the newborn. One of the blood spots that were collected for amino acid screening was used to estimate the amount of G6PD. The spots were cut out and eluted into buffer and electrophoresed. After electrophoresis, the enzyme was located on cellulose acetate membrane by activities staining. G6PD converts glucose-6-phosphate to 6-phosphogluconic acid and the reducing equivalent are transferred from NADP to NADPH. NADPH is nonenzymatically oxidized by phenazinemetosulphate which transfers the electrons to colourless nitrobluetetrazolium to yield insoluble blue pharmanzan derivatives (Grunbaum 1981). The intensity of the blue colour is marked 0, +, ++, + + +. The scoring was compared in a double blind manner with the values obtained on quantitative estimation of enzyme activity. The correlation of

scoring was excellent. Table 11 depicts data on G6PD deficiency in 2,980 subjects. The normal values range from 8 to 13IU, whereas values between 0-6 are classified as deficient. As mentioned earlier, in an X-linked autosomal disorder, the female is the carrier and the male children are affected. The data in table 11 indicates that there are as many female deficient as the male children—a paradoxical finding that needs to be explained. Many of the studies carried out by earlier investigators involved sample sizes ranging from 50-100 and in Jamnagar study it was a tribal population (Aluwalia 1984, Anand et al. 1981 and Pandya 1972). In a recent study in Saudi Arabia (Hazmi & Warsy 1986) 4000 individuals were examined for G6PD deficiency and sex ratio of 1.1 was observed. Several explanations can be offered for this observation. The easy one would be to say that somehow the process of inactivation of X-chromosome is interfered. But this explanation dose not provide a satisfactory molecular basis for the phenomenon. We had, however, preferred a hypothesis that an autosomal gene product was regulating the expression of G6PD gene located on the X-chromosome. A few months ago Yoshida and his associates (Hitoshi et al. 1989) showed that an N-terminal segment of the enzyme was coded by a gene sequence present on an autosome and the rest of the enzyme by the X-chromosome. If there is a mutation on the autosome one could explain the sex ratio of 1:1. For reasons as yet not clear either due to inbreeding or due to a Founder effect this mutation could have been concentrated in Saudi Arabia and India as historically there is connection between these two populations. At the present time this is a highly speculative concept drawing support from unconfirmed observations (Hitoshi et al. 1989). Alternative hypothesis cannot be ruled out.

**Table 11**  
*Screening of newborn children (2980) for G-6-PD deficiency*

| Sex    | Deficient | Normal |
|--------|-----------|--------|
| Male   | 82        | 1437   |
| Female | 83        | 1378   |

Normal: 8-13 IU

Deficient: 0-6 IU

**Table 12**  
*G6PD sex ratio in other studies*

| Place        | Deficient (%) |        |
|--------------|---------------|--------|
|              | Male          | Female |
| Chandigarh   | 7.1           | 5.7    |
| Jamnagar     | 22.6          | 14.1   |
| Saudi Arabia | 2.9           | 2.0    |
| Bangalore    | 4.0           | 4.1    |

Another genetic disorder we examined in a subgroup of the population in order to probe into the effects of inbreeding was pseudocholinesterase deficiency. It was reported that many Vysyas suffer from scolineapoena, i.e. they do not recover rapidly from succinylcholine administered prior to anesthesia. They need to be hyperventilated after surgery and be attached to support systems for very prolonged periods of time.

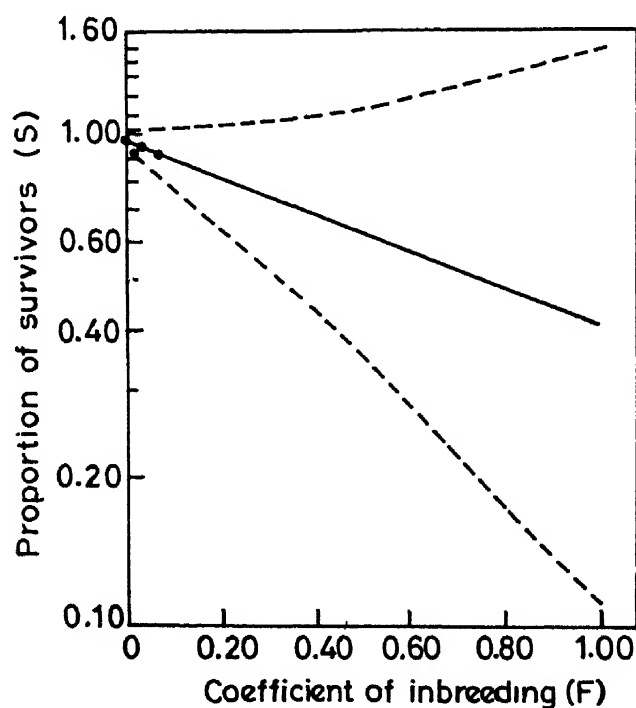


Figure 1

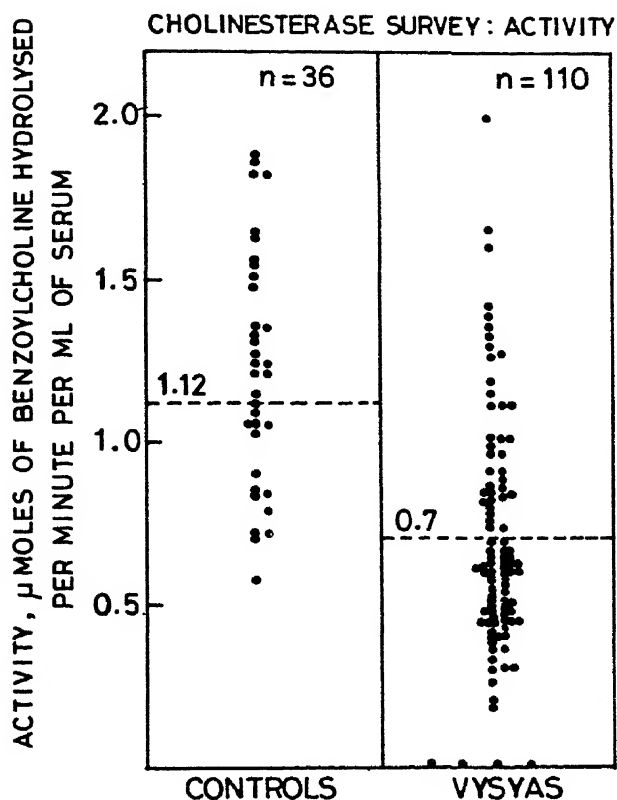


Figure 2

The enzyme activity was estimated as described earlier (Rao & Gopalan 1979). The enzyme is characterized by a change in the dibucaine and fluoride inhibition numbers. While the normal enzymes have fluoride and dibucaine numbers of 60 and 80, these values are changed for the mutant enzyme (Whittaker 1977).

The data on the activity levels in control in a small group of Vysyas is shown in figure 2. It is immediately apparent that the mean value in the Vysyas subjects is almost 50% of that of normals. In many of the individuals the activity is indeed very low. The dibucaine and fluoride numbers are not very different from the control values. These results would suggest that in Vysyas population we have a different type of mutation than is commonly encountered in the Western world. The mutation that was probably present could be of the silent type (Whittaker 1977). The most surprising aspect of this study was the observation that the Vysyas do not practice consanguinity. In fact in the study group there was only one consanguineously related offspring compared to about 33%

inbreeding this endogamous subgroup in an inbreeding milieu has concentrated at a higher level an autosomal recessive mutant gene. The simplest explanation could be that the founders of this group might have had the mutant gene which they could have passed on to the offsprings and the mutant allele could have reached equilibrium.

The studies that I have presented lead to the following conclusions. One is, that considering the population as a whole, inbreeding has *not* resulted in a large increase in the incidence of autosomal recessive disorders or marked decrease in fertility or increased infantile mortality. However, in preselected groups such as sick children or among the mentally retarded individuals, an effect of inbreeding is seen. It can therefore be concluded that rather than disturbing the established social practice of consanguinity, it would be advisable to provide genetic counselling to those families that have an affected child.

### ACKNOWLEDGEMENTS

This work was carried out in collaboration with Dr. A H Bittles, King's College, University of London, London UK; Dr A N Radhakrishnan, A Radha Rama Devi, H S Savithri, R Ramadevi, J K Acharya at the Indian Institute of Science, and Dr S Vankat Rao, at Central Food Technological Research Institute, Mysore. Ms Seetha S Murthy, Shantha Kumari, T N Veena, and Padmamma in addition to providing able technical assistance charmed more than 100,000 mothers to permit the collection of three drops of blood from the newborn infants, as well as answering several questions. Financial assistance from the Indian Institute of Science, Overseas Development Agency, and British Council Division of the UK Government, and the Indian Council of Medical Research is gratefully acknowledged.

### References

1. Ahluwalia A J Punjab Survey; *Res Bull Punjab Univ Sciences* **34** (1984) 153-160
2. Anand J S, Baxi A J, Nirmal R J, Adalja R and Anand A K A study of Glucose-6-Phosphate Dehydrogenase activity in Bhanushali and non-Bhanushali children; *Indian Paed* **18** (1981) 385-388
3. Appaji Rao N, Radha Rama Devi A, Savithri H S, Venkat Rao S and Bittles A H Neonatal screening for aminoacidaemias in Karnataka, South India; *Clin Genet* **34** (1988) 60-63



4. Bittles A H and Makov E Linear regressions in the calculations of lethal gene equivalents in man; *Ann Hum Biol* 12 (1985) 287-289  
 —, Radha Rama Devi A, Savithri H S, Rajeshwari Sridhar and Appaji Rao N 1985 Inbreeding and postnatal mortality in Karnataka, South India; *J Genet* 64 135-142  
 —, —, —, —, — and — Consanguineous marriage and postnatal mortality in Karnataka, South India; *Man. (N.S)* 22 (1987) 736-745
5. Cavalli-Sforza L L and Bodmer W F (Sanfrancisco: Freeman W H & Co) (1971) 341-476.
6. Darwin G Marriages between first cousins in England and their effects; *J Stat Society* 38 (1875) 153-184
7. Freire-Maia N *Eugen A* 15 (1968) 22
8. Garrod A E The incidence of Alkaptonuria: A study in chemical individuality; *Lancet* 2 (1902) 1616-1620
9. Grunbaum W *Hand Book of Forensic Medicine Individualization of Human Blood and Blood Stains Sartorius GmbH*; (1981) p 73
10. Hazmi M A F and Warsy A S Glucose-6-Phosphate Dehydrogenase polymorphism in the Saudi Population; *Hum Hered* 36 (1986) 24-30
11. Hitoshi K, Yin Hueng Y W and Yoshida A Two structural genes on different chromosomes are required for encoding the major subunit of human red cell Glucose-6-Phosphate Dehydrogenase; *Cell* 58 (1989) 595-606
12. Ireland J T and Read R A A thin layer chromatography methods for use in neonatal screening to detect excess aminoacidaemia; *Ann Clin Biochem* 9 (1972) 129-132
13. Luzzato L Glucose-6-Phosphate Dehydrogenase; *Advances in Human Genetics* 15 (1985) 217-303
14. Morton N E, Crow J F and Muller H J An estimate of the mutational damage in man from data on consanguineous marriages; *Proc Natl Acad Sci (USA)* 42 (1956) 855-863
15. Pandya SV Erythrocyte G6PD deficiency. Dissertation submitted to Saurashtra University (1972)
16. Radha Rama Devi A, Appaji Rao N and Bittles A H Consanguinity, fecundity and post natal mortality in Karnataka, South India; *Ann Hum Biol* 8 (1981) 469-472  
 —, — and — Inbreeding and the incidence of childhood genetic disorders in Karnataka, South India; *J Med Genet* 24 (1987) 362-368
17. Rao P R and Gopalan K B High incidence of the silent allele at cholinesterase locus 1 in Vysyas of Andhra Pradesh (S India); *Hum Genet* 52 (1979) 139-142

18. Rao PSS and Inbaraj S G Trends in human reproductive wastage in relation to n-long term practice of inbreeding; *Ann Hum Genet* **42** (1979) 401-413
19. Sanghvi L D Inbreeding in India; *Eugen Quart* **13** (1966) 291-301
20. Sridhara Rama Rao B S, Subhas M N and Narayanan H S Metabolic anomalies detected during a systematic biochemical screening of mentally retarded cases; *Indian J Med Res* **65** (1977) 241-247
21. Levels, trends and differentials in fertility (New Delhi, Ministry of Home Affairs:) Vital Statistics Division Office of the Registrar General, India (1982)
22. Whittaker M *Estimation of Plasma Cholinesterase Activity and the Use of Inhibitors for the Determination of Phenotypes* (1977)



**Pabitra Kumar Maitra** (b. 1 November 1932) did D.Phil. (1960) from University of Calcutta. He is Senior Professor, Tata Institute of Fundamental Research, Bombay; was previously Visiting Professor, University of Illinois, Chicago (1982-83), University College, Cardiff (1985).

Maitra established the genetics of the glycolytic pathway in yeast and bacteria. Discovered two new enzymes of glycolysis in yeast.

Maitra is Fellow of Indian Academy of Sciences. He is the recipient of J.C. Bose Medal (INSA) (1992).

*Pabitra Kumar Maitra was elected to the fellowship of the Academy in 1980*

**Address:** Professor, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay 400 005; Tel: (022) 2152971 (Ext 248); Telex: 011-83009 TIFR IN; Fax: 91-22-2152110; E-mail: maitra@tifrvax.bitnet; Res. 801, Bhaskara, TIFR Housing Colony, Homi Bhabha Road, Colaba, Bombay 400 005; Tel. (022) 2152499.

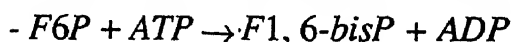
## ALLOSTERIC ENZYMES OF THE GLYCOLYTIC PATHWAY

P K MAITRA AND ZITA LOBO

*The paper describes mutations in the coding sequence of two glycolytic genes in the yeast *Saccharomyces cerevisiae*. These genes encode the structures of the allosteric enzymes fructose 6-phosphate kinase and pyruvate kinase respectively. The mutations bring about loss of allosteric control in these enzymes altering the substrate saturation profile from a sigmoidal to a hyperbolic one and abolition of heterotropic interaction with their respective modulators -ATP for phosphofructokinase and fructose 1, 6-bisphosphate for pyruvate kinase. The de-regulated mutant with a defective regulatory subunit in phosphofructokinase does not affect aerobic growth very substantially. The mutant with the non-allosteric pyruvate kinase was similarly competent for gluconeogenic growth.*

### INTRODUCTION

The glycolytic pathway operates in two directions, from glucose to alcohol or from alcohol to glucose. A common set of enzymes is used for both the directions except for two irreversible steps, -fructose 6-P kinase and pyruvate kinase. In the yeast *Saccharomyces cerevisiae* these two enzymes are strongly regulated not only by their respective substrates but also by metabolites which are not their biochemical neighbours. This mode of regulation, called allosteric control, is said to provide direction to metabolic flow by adjusting the concentration of the regulatory metabolites. Fructose 6-P kinase which catalyses the reaction.



displays sigmoidal saturation kinetics with fructose 6-P. Furthermore, ATP inhibits the enzyme strongly at low concentrations of fructose 6-P. A variety of metabolites which activate the enzyme allosterically do so by de-inhibiting it against ATP.

The other allosteric enzyme in the glycolytic pathway is the yeast pyruvate kinase, the reaction being



This enzyme also shows sigmoidal saturation with the substrate P-pyruvate, saturating at rather high concentrations which are unphysiological. The upstream glycolytic metabolite, fructose 1,6-bis P activates the enzyme in micromolar concentrations transforming the sigmoidal saturation profile to hyperbolic. It is believed that this positive control provides the switching mechanism that allows P-enol pyruvate to be either conserved for glycolytic reversal (Gancedo et al. 1967) or speedily utilised for fermentation.

We describe here mutations in the yeast *S. cerevisiae* which disturb the control properties of these two enzymes and examine the effect of this derangement on the physiology of the cell *in vivo*. Some of these results had been published earlier (Lobo & Maitra 1982, Maitra & Lobo 1977).

## MATERIALS AND METHODS

The wild-type strains of yeast and the mutants have been described in earlier publications (Nadkarni et al. 1984). Enzymes were assayed by continuous fluorescence measurements (Maitra & Lobo 1971). For inducing cytoplasmic petites, ethidium bromide was used as the mutagen and glucose plates for selection. Growth curves were constructed both by turbidity measurements at 650 nm as also by viable cell titre.

## RESULTS

### *Regulatory Features of Wild-type and Mutant Phosphofructokinases*

The sigmoid curve in figure 1 depicts a canonical profile of the velocity of an allosteric enzyme. Below the substrate concentration, say 4 in the figure, the velocity is very much lower than that for the hyperbolic curve, then rises rather abruptly around  $[S] = 7$ , and both tend to reach the same maximal velocity. Of particular interest from the view point of metabolic regulation is the observation that an allosteric modulator causes the sigmoid profile to change to hyperbolic with an attendant steep increase in reaction velocity without any increase in the substrate concentration (Monod et al. 1963). This would be illustrated here for two glycolytic enzymes, fructose 6-P kinase and pyruvate kinase.

Phosphofructokinase has the subunit composition  $\alpha_4\beta_4$  consisting of four of catalytic subunit  $\beta$  and four each of regulatory subunit  $\alpha$  encoded respectively by the genes *PFK1* and *PFK2* (see Gayatri & Maitra 1991). Mutations leading to extinction of the soluble phosphofructokinase activity map only to *PFK1*, while lesions that lead to altered regulation

map to *PFK2*. An enzyme from a strain *PFK1 pfk2* may either be insensitive or hypersensitive to ATP, depending on the particular mutant allele. The results in figure 2 depict the velocity of the enzymes from the wild-type and an ATP insensitive mutant as a function of the substrate concentration. The ATP profile shows how different the mutant enzyme is from the wild-type enzyme. The latter is highly susceptible to inhibition by small concentrations of ATP while the former has completely lost this sensitivity. The saturation profile of the velocity with the concentration of fructose 6-P mirrors the ATP behaviour. The wild-type enzyme shows a sigmoidal manner of saturation while the cell with *pfk2* mutation synthesizes a typically Michaelis enzyme. Fructose 2, 6-bisphosphate, the potent activator of the wild-type enzyme has no effect on this enzyme.

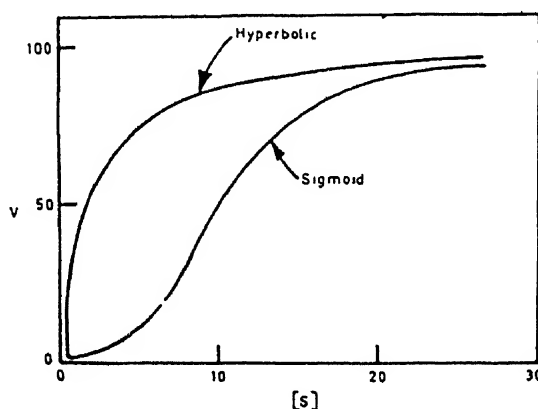


FIG 1 Hyperbolic and sigmoidal saturation profile of two enzymes: V stands for relative velocity, [S] for the concentration of the substrate in an arbitrary scale

The results in figure 3 provide a relative study of the saturation with fructose 6-P of three mutant enzymes (I), a non-allosteric allele *pfk2-5* (II), the wild-type *PFK2* and (III) a hyper-allosteric allele *pfk2-4* that is highly sensitive to inhibition by ATP. The apparent affinity for fructose 6-P at 0.07 mM ATP was very low for the non-allosteric enzyme (0.05 mM), and very high for the hyper-allosteric enzyme (4.5 mM). At the higher ATP concentration of 0.4 mM the apparent  $K_m$  for the non-allosteric enzyme was 0.12 mM fructose 6-P, while the hyper-allosteric enzyme was totally inhibited. The non-allosteric enzyme is much less sensitive to ATP inhibition than the hyper-allosteric enzyme; the latter is also stimulated strongly by fructose 2, 6-bisphosphate.

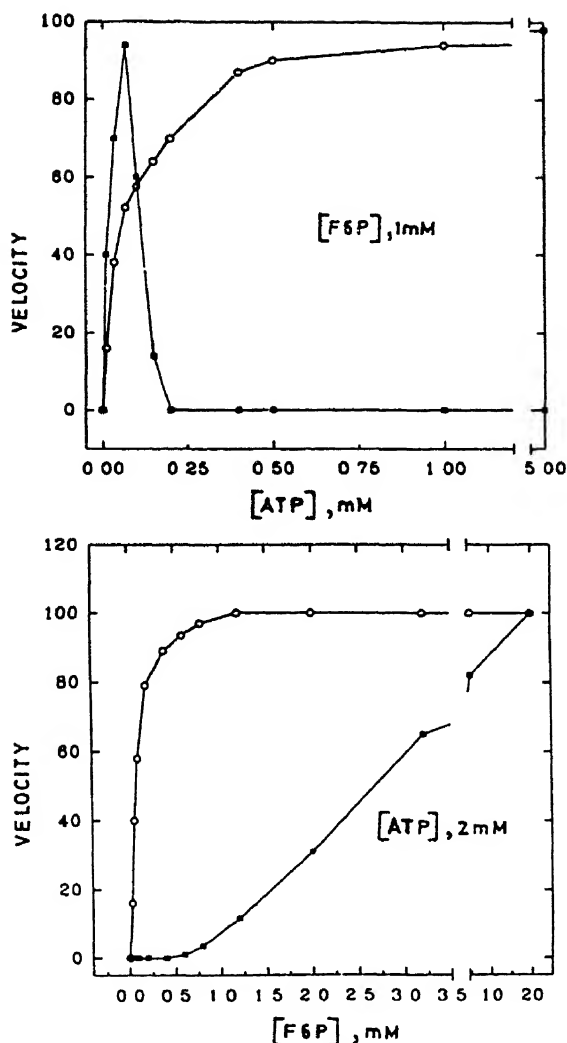


FIG 2 Velocity of yeast phosphofructokinase as a function of substrate concentration: The curves with filled squares refer to the wild-type enzymes and the lines with unfilled circles to a non-allosteric mutant. Partially purified preparations of the enzymes from the respective strains were used. The concentrations of the fixed substrates are also indicated.

### *A De-regulated Pyruvate Kinase*

During the growth of yeast on non-fermentable carbon sources such as ethanol, synthesis of P-enol pyruvate from oxalacetic acid is needed as also a safeguard against its metabolic slide to pyruvate. This can be achieved if the levels of fructose 1, 6-bisphosphate are low so that pyruvate kinase activation by this metabolite is prevented and P-enol pyruvate conserved for a gluconeogenic flux (Maitra & Lobo 1978) rather than a glycolytic short-circuit to pyruvic acid. A de-regulated pyruvate

kinase escaping this control might jeopardise glycolytic reversal and therefore growth on alcohol.

Starting with a structurally altered mutant *PK3* of pyruvate kinase unable to grow on glucose, a spontaneous revertant *PK3R121* was isolated that regained glucose-positivity and synthesized a variant enzyme completely independent of fructose 1, 6-bisphosphate for its catalytic activity. Results in figure 4 indicate the profile of saturation of this enzyme with the substrate P-enol pyruvate. The result shows that the mutant enzyme saturates in a hyperbolic manner, however, the wild-type enzyme displays a pronounced sigmoidicity. Addition of small concentrations of fructose 1, 6-bisphosphate transforms the wild-type enzyme to a form virtually identical with the mutant enzyme, while the latter is without effect. The mutation thus seems to have frozen the pyruvate kinase in a conformation comparable to that attained by the wild-type enzyme in the presence of the activator fructose 1, 6-bisphosphate.

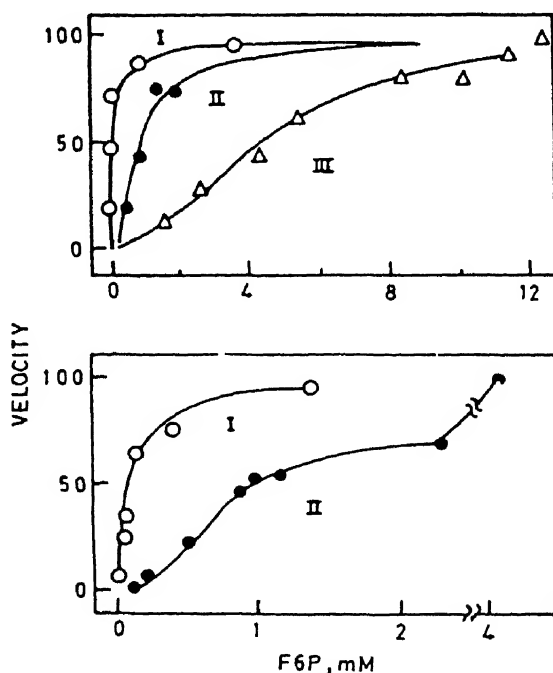


FIG 3 Dependence of phosphofructokinase activity on the alleles of the regulatory gene *PFK2* as influenced by the concentration of fructose 6-P. *PFK2*, *pfk2-5* and *pfk2-4* represent the wild-type, a non-allosteric and a hyper-allosteric alleles respectively. The fixed ATP concentrations in the upper and lower panels are 0.07 and 0.4 mM respectively (M. Gautam, Ph.D. thesis to Bombay University, 1985).



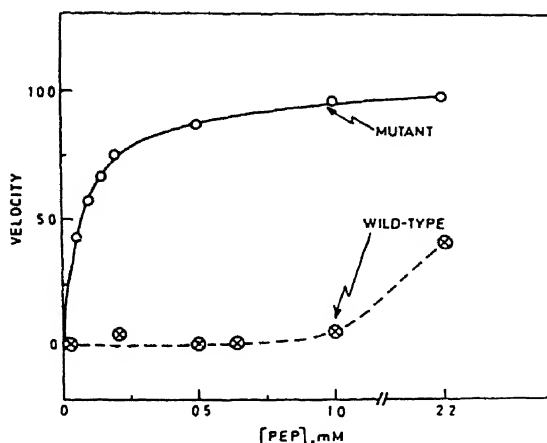


FIG 4 Reaction velocity profile of yeast pyruvate kinase as a function of the concentration of P-enol pyruvate. The continuous line refers to a partially purified enzyme from the mutant PK3R121 and the dashed curve to that from a wild-type strain. The velocity is in arbitrary units

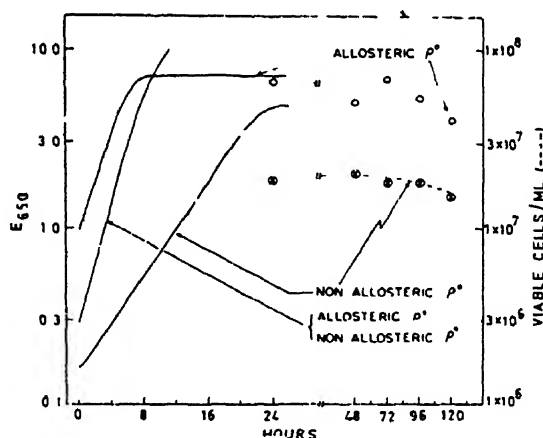


FIG 5 Growth curves of mutants with non-allosteric phosphofructokinase in yeast extract-peptone-glucose medium. Strains used were: wild-type cytoplasmic petite (allosteric,  $p^0$ ); *PFK1-pfk2-5*, cytoplasmic petite (non-allosteric,  $p^0$ ); wild-type, respiratory-competent (allosteric,  $p^+$ ), *PFK1 pfk2-5*, respiratory-competent (non-allosteric,  $p^+$ ). The dashed curves refer to viable cell count

### Growth Properties of De-regulated Mutants

Figures 5 and 6 summarise the results of the experiments on growth kinetics of the mutants lacking the regulatory properties of phosphofructokinase and pyruvate kinase respectively. The mutants in phosphofructokinase have also been examined for the effect on anaerobic growth by rendering them cytoplasmic petites ( $p^0$ ) with ethidium bromide as anaerobic experiments are harder to control. Results in figure 5 show

that respiratory-competent ( $\rho^+$ ) non-allosteric mutants of phosphofructokinase grow as fast as the allosteric  $\rho^+$  parent. During non-respiring conditions, however, non-allosteric mutants grow somewhat more slowly than their allosteric counterparts. It should be pointed out here that such results have not been confirmed yet for an isogenic pair of strains.

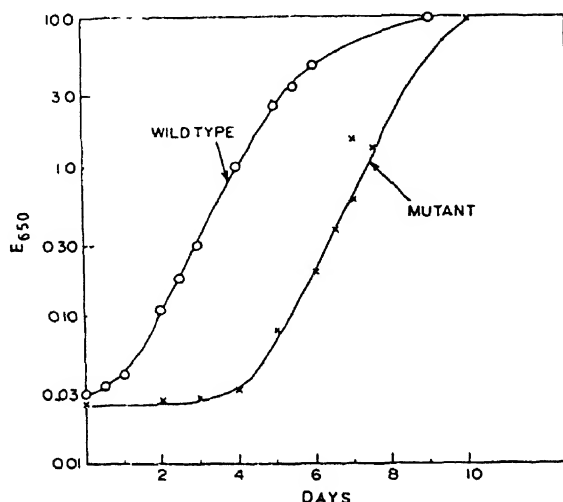


FIG 6 Kinetics of growth of a wild-type yeast strain and a mutant yeast synthesizing an FDP-insensitive pyruvate kinase in a minimal medium containing ethanol as a carbon source

How does the physiology of the cell get affected by the loss of allosteric control on yeast pyruvate kinase? The mutant with the fructose 1, 6-bisphosphate-insensitive pyruvate kinase grows on a minimal salts-alcohol medium with much the same facility as does the wild-type strain. The main difference between these two strains lay in the longer delay in the mutant in the onset of growth than in the wild-type strain. Once growth commenced, the rates were nearly the same. This showed that loss of fructose 1, 6-bisphosphate regulation on pyruvate kinase does not substantially affect the growth behaviour of yeast cells either on glucose or on alcohol.

## DISCUSSION

The experiments described here relate to the control characteristics of two enzymes of the glycolytic pathway, fructose 6-P kinase and pyruvate kinase. Both of these enzymes are said to regulate two irreversible steps of the pathway. This regulation has come about by evolutionary acquisition

of special structural features in these proteins such as creation of a distinct allosteric site either in the same subunit (as in pyruvate kinase) or in a different subunit, as in fructose 6-P kinase, Binding of the modulator molecules to these sites brings about changes at the catalytic site that lead to appropriate response to the reaction velocity.

How do the allosteric properties of pyruvate kinase or phosphofructokinase serve the needs of the cell ? Is their primary objective the control of gluconeogenesis or glycolysis ? The fact that the mutant with pyruvate kinase insensitive to fructose 1, 6-bisphosphate is able to grow on alcohol and maintain significantly high levels of P-enol pyruvate showed that control by the bisphosphate could not be the sole mechanism *in vivo* to control gluconeogenesis. Likewise our failure to observe any substantial derangement of growth either on alcohol or on glucose in the non-allosteric phosphofructokinase-bearing strain points to the same conclusion that our experimental design may not have been subtle enough. Robinson and Fraenkel (1978) had reached a similar conclusion on the relevance of allostery of *E. coli* phosphofructokinase as a determinant of growth.

## References

1. Gancedo C, J M, Gancedo and Sols A Regulation of the concentration or activity of pyruvate kinase in yeasts and its relationship to gluconeogenesis, *Bio-chem. J.* **102** (1967) 23C-25C
2. Gayatri A G and Maitra P K Complementation of soluble phosphofructokinase activity in yeast mutants; *Eur. J. Biochem.* **200** (1991) 149-155
3. Lobo Z and Maitra P K Genetic evidence for distinct catalytic and regulatory subunits in yeast phosphofructokinase; *FEBS Letters* **139** (1982) 93-96
4. Maitra P K and Lobo Z Reversal of glycolysis in yeast; *Arch. Biochem. Biophys.* **185** (1978) 535-543  
— Yeast pyruvate kinase; A mutant form catalytically insensitive to fructose 1, 6-bisphosphate; *Eur. J. Biochem.* **78** (1977) 353-360  
— — A kinetic study of glycolytic enzyme synthesis in yeast; *J. Biol. Chem.* **246** (1971) 475-488
5. Nadkarni M, Parmar L, Lobo Z and Maitra P K Mutations in the regulatory subunit of soluble phosphofructokinase from yeast; *FEBS Letters* **175** (1984) 294-298
6. Monod J, Changeux J P and Jacob F Allosteric proteins and cellular control systems; *J. Mol. Biol.* **6** (1963) 306-329



**Pran Nath Mehra** (b. 27 October 1907) did D.Sc. (1942) from Panjab University, Lahore. He is Research Advisor, Himachal Pradesh Agricultural University and Professor Emeritus (currently), Punjab University, Chandigarh.

Mehra put forth a new theory, known as 'condensation theory', on the origin and evolution in Hepaticae, which has been widely accepted all over the world. He has done extensive work on: phyletic classification of ferns; cytogenetic evolution in hardwoods based upon investigations on more than 500 species of trees in the Himalayas,

considered to be the most comprehensive investigation of forest trees in the world; and induced meiotic reductions in root-tips. Also put forth a saturation hypothesis for the absence of polyploidy in conifers, and gene block hypothesis as applied to ferns and flowering plants. Other areas in which Mehra has made contributions are phyletic evolution of grasses, and distribution of Podocarpean flora and the concept of Gondwanaland. He is the Author of 9 research monographs, reviewed by specialists as very significant contributions; served as member of editorial board of *The Wealth of India* (CSIR).

Mehra is Member of National Academy of Sciences (India), Indian Botanical Society, Indian Palynological Society, Indian Society of Tree Scientists, and Indian Society of Pharmacognosy; Corresponding Member, Botanical Society of America; He was Member, INSA Council (1969-70). He is the recipient of Birbal Sahni Gold Medal (1968); UGC National Lecturer (1970); International Botanical Congress Medal as Vice President (1975); Gunner Erdtman International Gold Medal (1975); Sunder Lal Hora Medal (INSA) (1984); Tree Scientist Gold Medal (1984); Birbal Sahni Centenary Commemoration Lecture Award and Gold Medal (1992); Padma Shri (1972).

*Pran Nath Mehra was elected to the fellowship of the Academy in 1952.*

**Address:** Bungalow 1055, Sector 27-B, Chandigarh 160 027; Tel: (0172) 26791.

## **SOME ASPECTS OF DIFFERENTIATION IN FLOWERING PLANTS**

P N MEHRA

### **INTRODUCTION**

If one observes an individual plant of pansy one is struck with the identical shape and symmetry of its leaves, their phyllotactic arrangement, and the size, architecture and similar colourful patterns of its flowers and their individual parts like petals, sepals etc. The same marvellous situation holds in a gladiolus, an orchid, or a lily, in which flower after flower possesses identical shapes and symmetry in the very minutest details. The same patterning of flowers in capitula in a *Chrysanthemum* is noticeable with infallible regularity. In a coleus plant identical serration of margin, similar colour pattern and shades on lamina are manifested in every leaf of an individual, which indeed is remarkable. It is obvious that there is something of a regular programming in the building up of the tissues constituting the various organs and identical functioning of their cells to produce the same architecture and colourful patterning. How does this all happen? What is the underlying basis for this development? This is now visualised to be the expression of various processes imprinted in an organism in the form of genes whose precise functioning in slabs and blocks, simultaneously or successively, leads to the organisation of similar specific tissues and organs through the synthesis of a multiple variety of proteins and other metabolic substances and their aggregation to form tissues and organs.

In recent years a transparent nematode, *Caenorhabditis elegans*, which is rather small in size, barely 1 mm long, has attracted the attention of a number of investigators who have meticulously followed its development from the fertilized egg to the adult stage in the living organism. This has been possible because it is easy enough, with suitable sophisticated type of microscope and optics, to follow the divisions of the cells at live stage and their organisation to form tissues in the adults. This organism is ideal in the sense that it completes its full life-cycle in only 3 days from conception to maturity. The fertilised egg undergoes successive invariant cell divisions to form precisely 945 cells in every individual of a progeny. These cells organise themselves into two blocks: one block of

302 cells (originally 407 precursor cells of which the remaining undergo programmed cell death) are earmarked as neural cells or neurons and the remaining 643 are the non-neural cells responsible for the formation of rest of the body parts. These latter undergo independent polyclonal divisions to form a mosaic of tissue composed of groups of cells derived from the same cell-lineage from which are built up all the other body parts, except for a lone specialised cell called the gonadal anchor cell which is set apart by the mutual interaction of the adjacent cells and has been activated to perform the special function of transforming a group of ventral hypodermal cells into a vulva. If the anchor cell is destroyed by a laser beam, no vulva is formed. Alternatively, if the other gonadal cells are destroyed excepting the anchor cell a vulva does form. Obviously the anchor cell is the key cell to direct the organisation of a vulva and it has developed to this status by the signalling and interaction of the other adjacent non-neural cells.

Similar wonderfully systematic programming is found to occur in the thoroughly worked out development of fertilised egg of fruit-fly or *Drosophila melanogaster*. The zygotic nucleus undergoes 9 successive divisions to produce 512 free nuclei which move to the periphery of the egg where 4 more divisions occur, leading ultimately to the organisation of a cellular blastoderm. At this stage a few pole cells are set apart in the posterior regions which are precursors of the germ line. Infolding of the blastoderm next occurs to form a segmented larva which is the result of action of at least 15 genes. This is the instar stage I, and is followed by instar stages II and III, and thereafter the timing for the larva comes to undergo metamorphosis to form a pupa as a result of the changing hormonal levels. This is also the time when certain proliferating cells are differentiated, called imaginal discs, which are 19 in number: of these 9 are paired (each pair committed to a particular developmental pathway to form left and right legs and left and right eye) and a fused genetical disc from which arise the genital organs.

Likewise *programmed development* seems to be occurring in plants as illustrated in the examples cited in the beginning. In simplistic terms, however, a plant when developing from a seed forms a block of active meristematic cells that constitute the stem apex along with two cotyledons at one pole, and a root primordium at the other end. When the stem grows further it sets apart blocks of cells which function to form leaves at regular intervals. At a later stage when nearing maturity, the stem apex transforms itself into a floral shoot, when blocks of cells are differentiated

successively in which floral genes become activated resulting in the formation of flowers.

Plant morphogenesis is much simpler than that of animals. It is not condensed and packaged as in the latter, but is protracted, recurrent and spread over space and time, although following the same basic principles of organisation. Much of the understanding of plant morphogenesis is facilitated by the discovery of plant hormones, and the ease with which callus tissues are raised from all of its parts *in vitro* which respond to the action of plant hormones.

In the following pages an attempt has been made to the understanding of plant morphogenesis from the *in vitro* cultures as well as from what occurs in nature in the fern and flowering plant systems.

### THE FERN SYSTEM

In a previous communication the author (Mehra 1975) had dealt with some aspects of differentiation in cryptogams. This included mainly the ferns, with some observations on mosses and liverworts. In ferns ten species at different ploidy levels (2x to 6x) were experimented upon which led to certain important generalisations in the Fern System as a whole. It was possible to invoke the gametophyte or the sporophyte from either generation, or their calli, at will using different concentrations of sucrose in the culture medium. The Fern System is rather simple in that it does not require complex morphogenetic substances for invoking transformations. On the basis of these experiments Gene-block Hypothesis was postulated which in essence meant that there were four Gene-blocks, each determining root, stem (rhizome), leaf, and the gametophyte. In each case the master gene initiated the activity first with the subordinate genes expressing themselves next in a sequential order resulting in the expression of an organ or a different generation. (In Ferns there are two different independent alternating generations in the normal life-cycle of the plant). It was also emphasised that it is not that easy to invoke the activity of subordinate genes responsible for differentiation of tissues bypassing the master gene. For example antheridia and archegonia must appear on the gametophyte and never on leaf, nor the mesophyll cells ever appear on gametophyte in the normal fern life-cycle.

In order to recapitulate some of the salient features of experimental morphogenesis, two ferns, *Pteris vittata* and *Cyclosorus dentata*, are dealt with here out of those investigated, for illustration.

*Petris vittata* is a tetraploid with  $n = 58$  chromosomes. The normal gametophyte is hence diploid. It can experimentally be made to develop directly diploid sporophyte by culturing it on Knudson's or any other suitable medium, with the prerequisite that it should contain 0.5-1% sucrose. In case such a sporophyte is brought back to a medium lacking sucrose, its parts can revert back to form the gametophyte again. Fig. 1B shows such reversion of upper pinnae of the apogamously produced leaf. It is equally interesting to notice the transformation occurring from typical sporophytic tissue, which possesses wavy walls of the epidermis interspersed with stomata and vascular tissue, to typical gametophytic cells with straight walls in the adjacent region (Fig. 1C, D). The switch for the expression of gametophyte has now been turned on leading ultimately to the expression of a typical gametophyte. Such interconversion is also noticeable in Fig. 1E where, from the gametophyte, sporophyte leaves have developed in response to sugar and are reverting back again to the gametophyte on culturing in the medium lacking sugar. This conversion is not confined to leaves only but even the root of apogamously developed sporophyte can grow into gametophyte directly (Fig. 1A, 2A).

In another experiment callus was derived from gametophytes by culturing them on Knudson's medium containing 2% sucrose or appropriate quantity of 2, 4-D. This gametophytic callus mass gave rise to a gametophyte when grown on Knudson's basal medium *without sugar* as expected (Fig. 2B), but when grown on the same medium containing 0.5% sucrose it exhibited a tendency to develop sporophytic tissues. This concentration of sugar seems to be at the threshold value as shown by some of its part developing typical gametophytic cells, while the other developing cells with wavy walls interspersed with stomata characteristic of the sporophytic generation (Fig. 2C-D). Such interconversion by the simple presence of sucrose is indeed amazing.

*Cyclosorus dentatus* too is a tetraploid with  $n = 72$  chromosomes. Calli were obtained from its root by culturing it on Knudson's basal medium supplemented with 1-4% sucrose and 0.5-4.0 mg/12,4-D and single callus cells were obtained with the help of shaker. Such single cells cultured on Knudson's basal medium *without sucrose* germinated like spores to form at first a filament which expanded later to form a heart-shaped gametophyte (Figs. 3A-C) which ultimately developed a cushion bearing antheridia and archegonia. These gametophytes unlike the ones produced in the normal life cycle were thus of 4x constitution rather than 2x. If, however, these cells were cultured on BM+ 0.5 sucrose they behaved



differently. Some of these formed flattened gametophytes which soon thickened at the notch region and produced scales characteristic of the sporophyte (Fig. 3D). In others a vascular strand, unusual for a gametophyte but characteristic of sporophyte, could be seen (Fig. 3E, F). Still others germinated to form a filament followed by expansion to form a strap-shaped thallus, which simply grew out to form the first leaf (Fig. 4A).

Strikingly, occasional unusual cases of the formation of sporangia on the gametophytes were also noticed. These are illustrated in Fig. 4B and on enlarged scale in Fig. 4C. In these it appears that the subset of the leaf Gene-block meant for the organisation of sporangia has been activated out of place in the gametophyte itself (normally it is in the leaf). Such cases, however, are very rare. In fig 4D is shown a gametophyte which has emerged from a surface cell of petiole of leaf cultured on BM without sugar.

Such interconversions are not confined to tetraploids only but have been experimentally induced in diploid ferns too, like *Ampetopteris prolifera* with  $n = 36$  and others (cf. Mehra 1975).

It is thus obvious that the Fern System is highly plastic and can be manipulated at will by simple experimental techniques. This is graphically represented in Fig. 5 where the entire genome is represented by a straight line for the sake of simplicity and is compartmentalised into four Geneblocks (see Mehra 1975 for details).

Our experience on *in vitro* differentiation in flowering plants has shown that morphogenetic development in these is similarly compartmentalised into concrete Gene-blocks. The experiments on a variety of herbs and trees form the basis of the present theme. These are illustrated species-wise in the following photographs. It must however be stated at the very outset that the manipulative substances is not just sugar but is much more complex, in the form of hormones—auxins and kinins, and natural growth regulators like coconut-water, yeast, malt extract etc.



FIG 1. *Pteris vittata* ( $n = 58$ , tetraploid)

A. Apogamous 2x root which has formed a gametophyte directly on basal Knudson's medium B. Normal 2x gametophyte cultured on Knudson's basal medium +0.5-2% sucrose produced an apogamous sporophytic compound leaf. When transferred to the same medium without sucrose the upper pinnae have been retransformed to gametophytes C. A portion of the upper region of the rachis. D. An islet of the transformed pinna highly magnified to show tracheids at arrow and large typical gametophytic cells below. E. Another gametophyte cultured on the same medium as in B showing transformation of the lamina of an apogamously developed leaf into a gametophyte

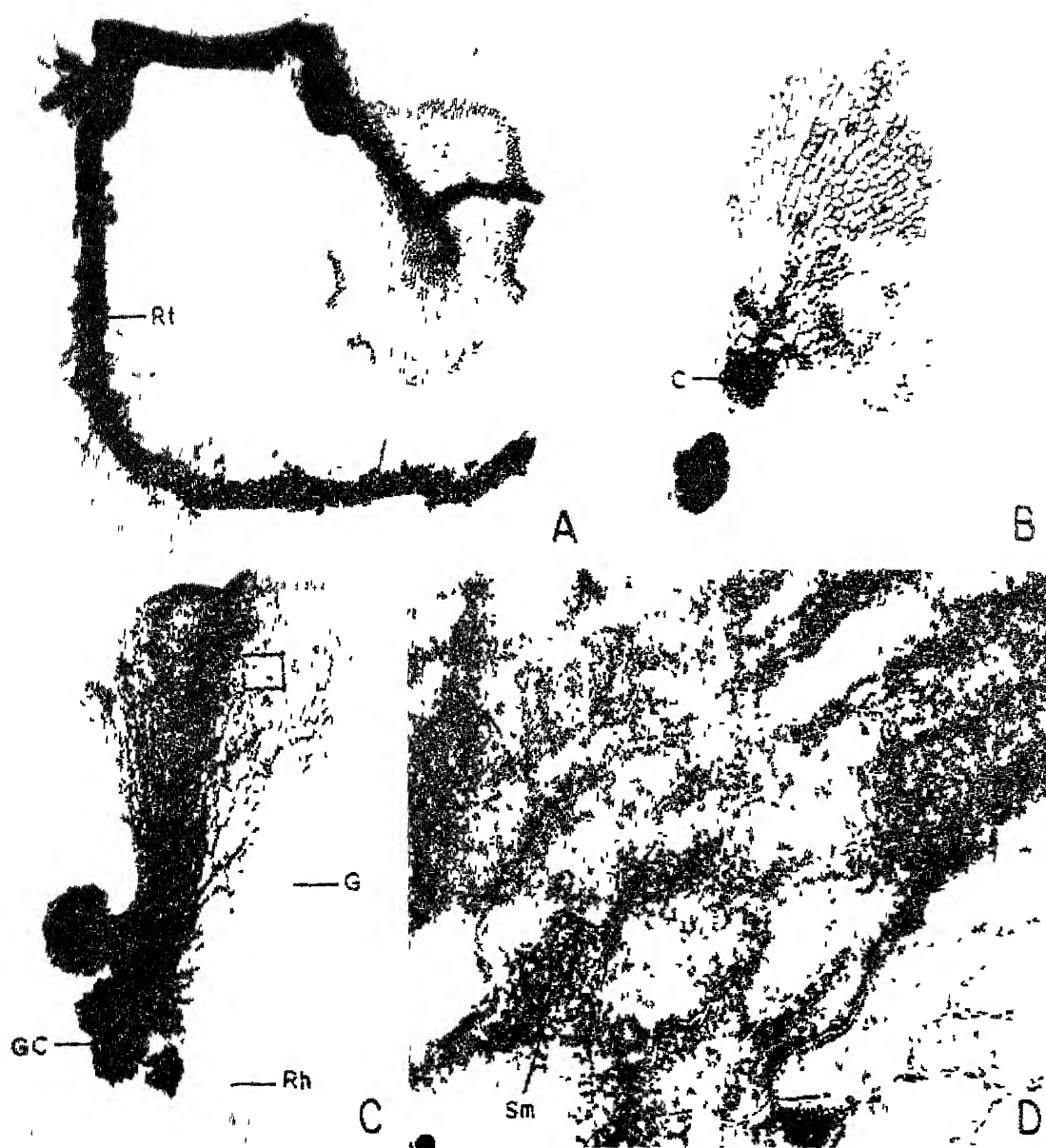


FIG 2. *Pteris vittata* contd

A. Excised root from diploid apogamous sporophyte which has grown out into a gametophyte on basal Knudson's medium B. Gametophytic callus mass forming a gametophyte on Knudson's basal medium C. Gametophytic callus mass on basal Knudson's medium containing 0.5% sucrose developing synthetic structures, part gametophytic and part sporophytic. D A magnified islet region of the above showing stomata and characteristic wavy walls of the leaf epidermal cells above and straight-walled gametophytic cells in the lower portion.

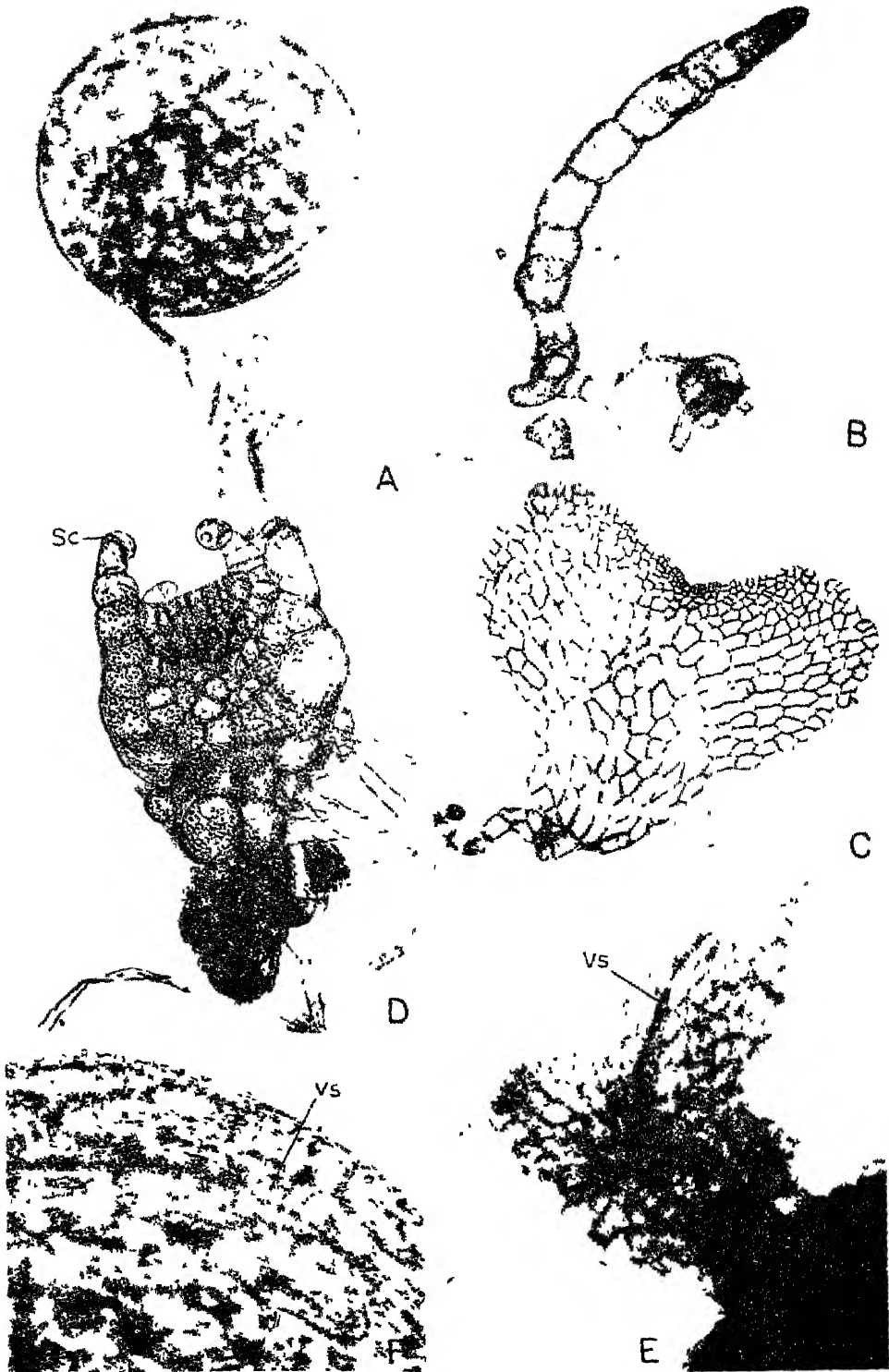


FIG 3 *Cyclosorus dentatus* ( $n = 72$ , tetraploid)

A. Root callus cell germinating to form a rhizoid on Knudson's basal medium. B. the same having produced a gametophytic filament. C. Further growth of the same to form a heart-shaped gametophyte on Knudson's basal medium. D. Root callus cell grown on Knudson's basal medium containing 0.5% sucrose. Note the activity at the apex of the gametophyte and formation of multicellular hairs characteristics of sporophyte. E. Formation of a vascular strand in the gametophyte cultured on the same medium as above. F. Vascular strand region of the above magnified.



FIG 4. *Cyclosorus dentatus* contd.

A. Root callus cell after forming a strap-shaped thallus has switched on to form a leaf from the anterior region when cultured on Knudson's basal medium containing 0.5% sucrose. B. Gametophyte at the apical region has produced young sporangia when cultured on Knudson's basal medium containing sucrose C. Magnified view of the sporangia. D. Leaf-stalk having developed a typical gametophyte.

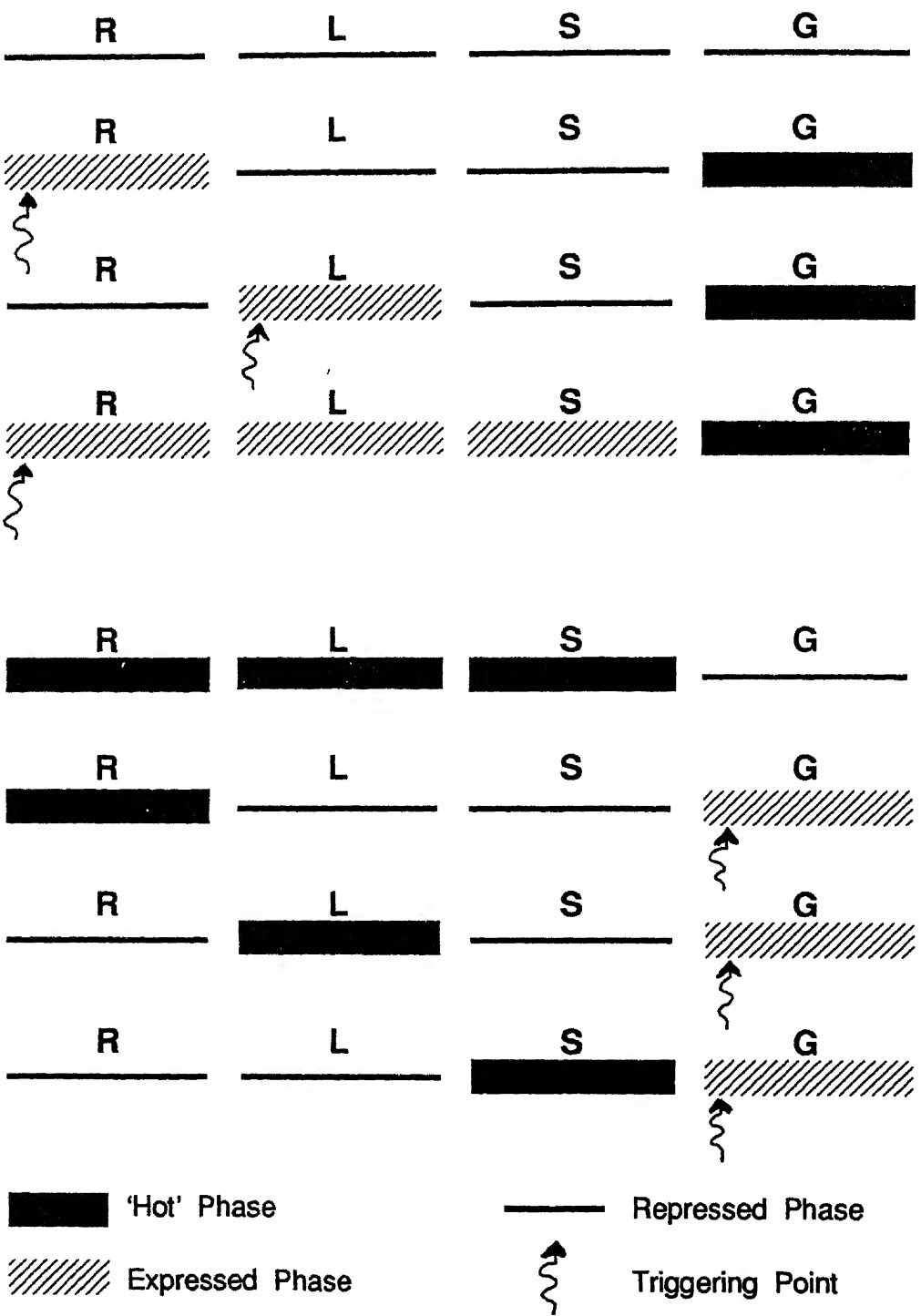


Fig 5 Gene-block hypothesis as applied to Fern System (Diagrammatic)

## GENE-BLOCK HYPOTHESIS AS APPLIED TO FLOWERING PLANTS

I believe that Gene-block hypothesis is equally applicable to flowering plants as stated above. Four Gene-blocks are visualised which seem to be operative in this group of plants during morphogenesis: one each for root, leaf, stem and flower. There are enough experimental as well as circumstantial evidences from nature to support this view point.

### EXPERIMENTAL EVIDENCES

I. The calli derived from explants of different plant parts such as root, leaf (also cotyledon), stem or flowers and their constituents can be made to differentiate either:

|                            |  |
|----------------------------|--|
| Roots only (Rhizogenesis)  | <i>Prunus amygdalis</i> (Fig. 6A, B)<br><i>Pterthea falconeri</i> (Fig. 12A)<br><i>Pyrus comuns</i> (Fig. 10A).  |
| Leaves only (Phylogenesis) | <i>Prunus amygdalis</i> (Fig. 6C)<br><i>Punica granatum</i> (Fig. 7A)<br><i>Pyrus communis</i> (Fig 10B)<br><i>Pterotheca falconeri</i> (Fig. 12B)   |
| Shoots (Caulogenesis)      | <i>Punica granatum</i> (Fig. 7B)<br><i>Malus pumila</i> (Fig. 9A, B))<br><i>Pyrus communis</i> (Mehra and Jaidka 1979)<br><i>Populus alba</i> (Mehra and Cheema 1980)<br><i>Pterotheca falconeri</i> (Fig 12C) |
| Flowers (Flourogenesis)    | <i>Punica granatum</i> (Fig. 7C)<br><i>Nicotiana</i> (Agheon 1962, 1965)<br><i>Chucory</i> (Paulet and Nitsch 1964)  |
| Or Embryoids               | <i>Punica granatum</i> (Fig. 8A-D)<br><i>Malus pumila</i> (Fig. 9C)<br><i>Pyrus comuns</i> (Fig. 10C, D)<br><i>Pterotheca falconeri</i> (Fig. 12D, E)  |

Such inductions have been invoked by incorporating plant growth hormones i.e. auxins, cytokinins and gibberellins, in various concentrations and combinations into the medium as indicated in the

captions to each figure. These inductions are depicted diagrammatically in Fig 13.

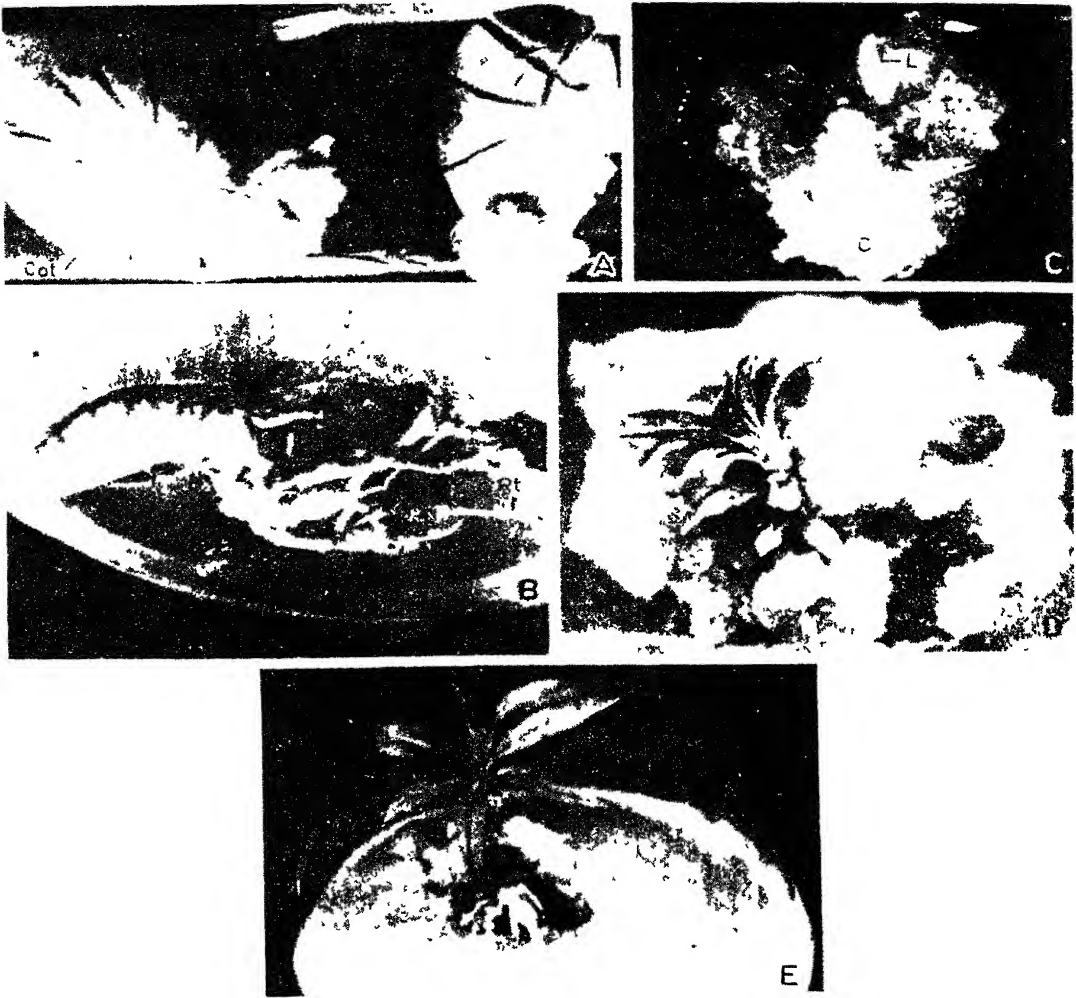


FIG 6. Almond (*Prunus amygdalis*)

A,B. Cotyledon callusing at the base and initiating roots on BMS + NAA (5 ppm) + CH (1 g/liter) C. Leaf callus differentiating isolated leaves on BMS + IAA (5 ppm) + CH (1g/liter). D. Leaf callus differentiating shoot on BMS + NAA (5 ppm) + CH (1g/liter) + K (1 ppm) E. Leaf callus differentiating whole plant on BMS + NAA (5 ppm) + CH (1g/litre) + K (1 ppm)



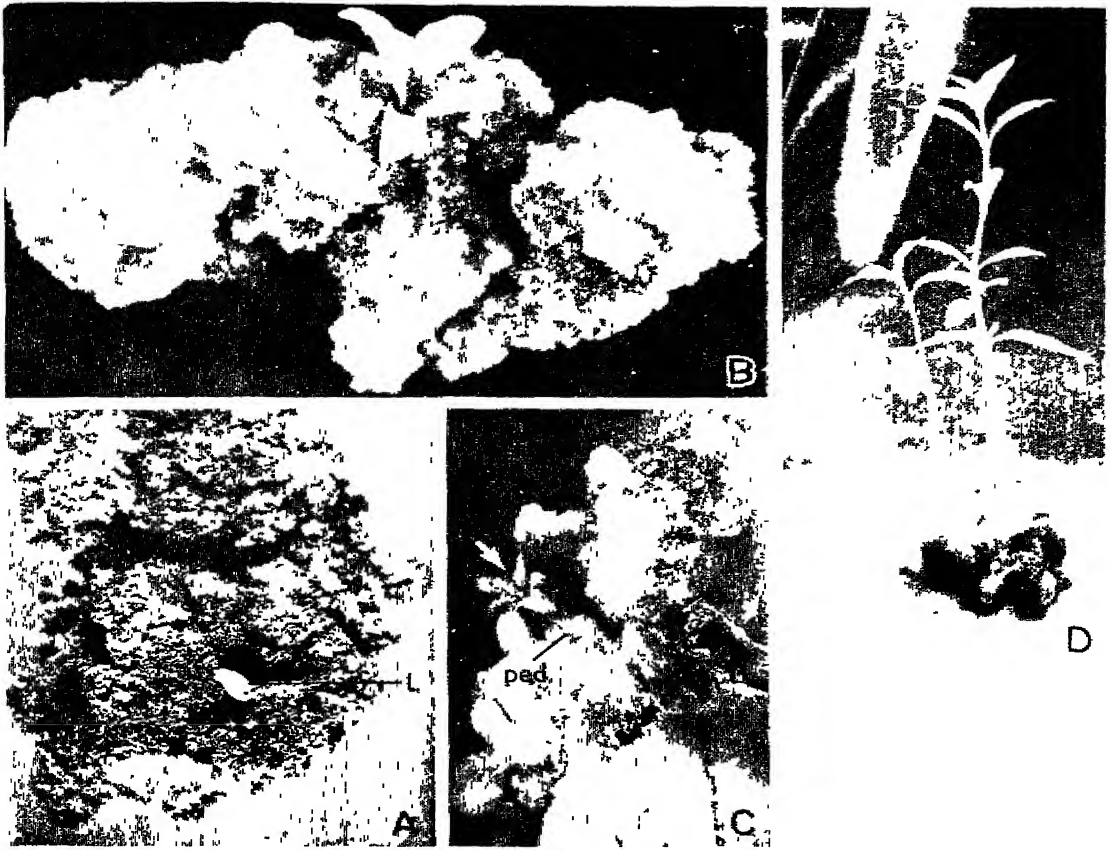


FIG 7. A-D Pomegranate (*Punica granatum*)

A. Stem callus on BMS + NAA (2 ppm) + BAP (2 ppm) differentiating isolated leaf B. Leaf callus on BMS + NAA (2 ppm) + BAP (2ppm) differentiating shoot. C. Leaf callus on BMS + NAA (2 ppm) + BAP (2 ppm) differentiating a flower (at arrow) with pedicel D. Cotyledon explant on BMS + NAA (2 ppm) + BAP (2 ppm) regenerating plants directly

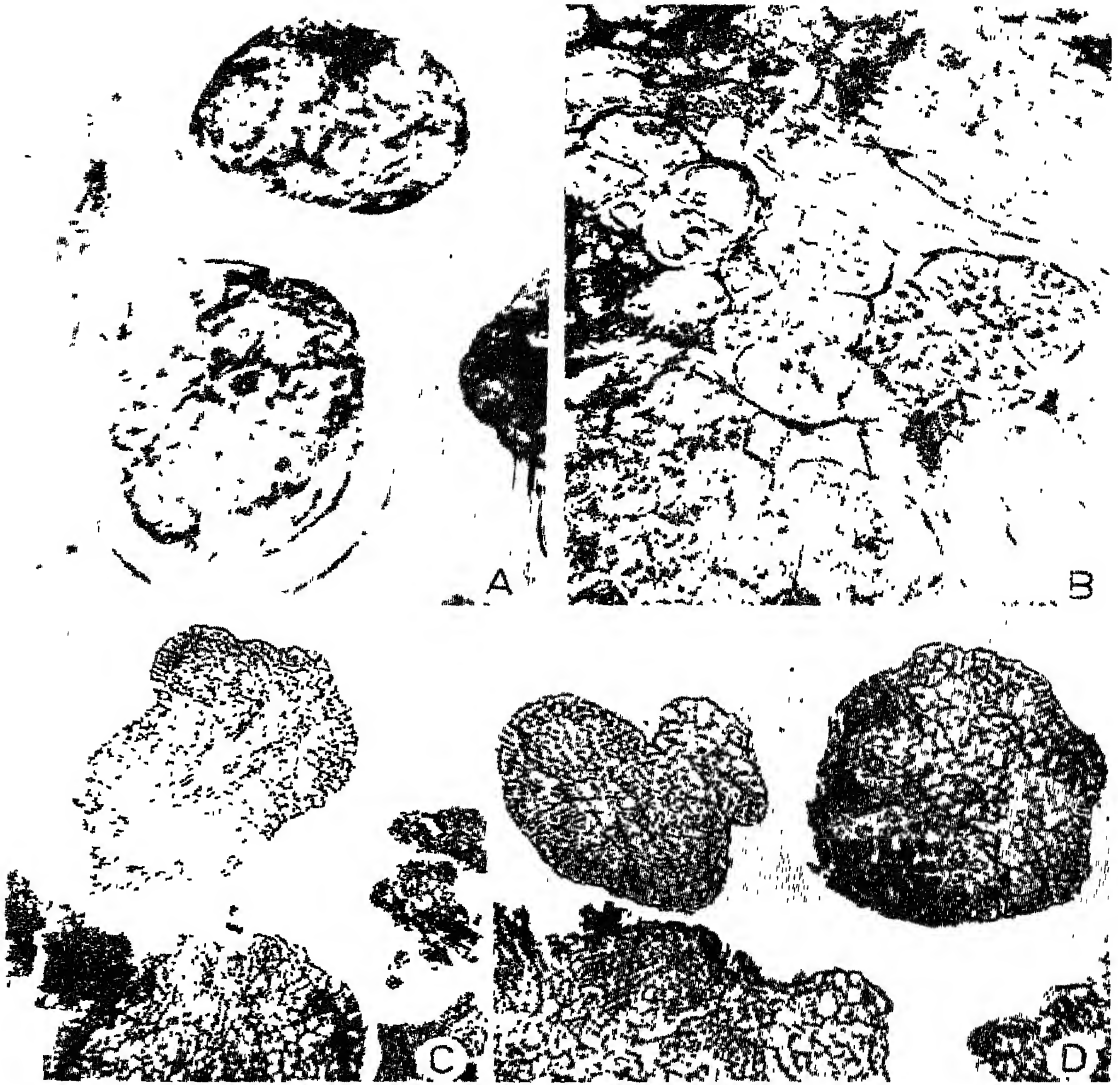


FIG 8 A-D Pomegranate contd Showing various stages of embryoid development A Two embryoid initials, note the gelatinous sheath around each B Embryoids at different stages of development C D. Fully developed embryoids.

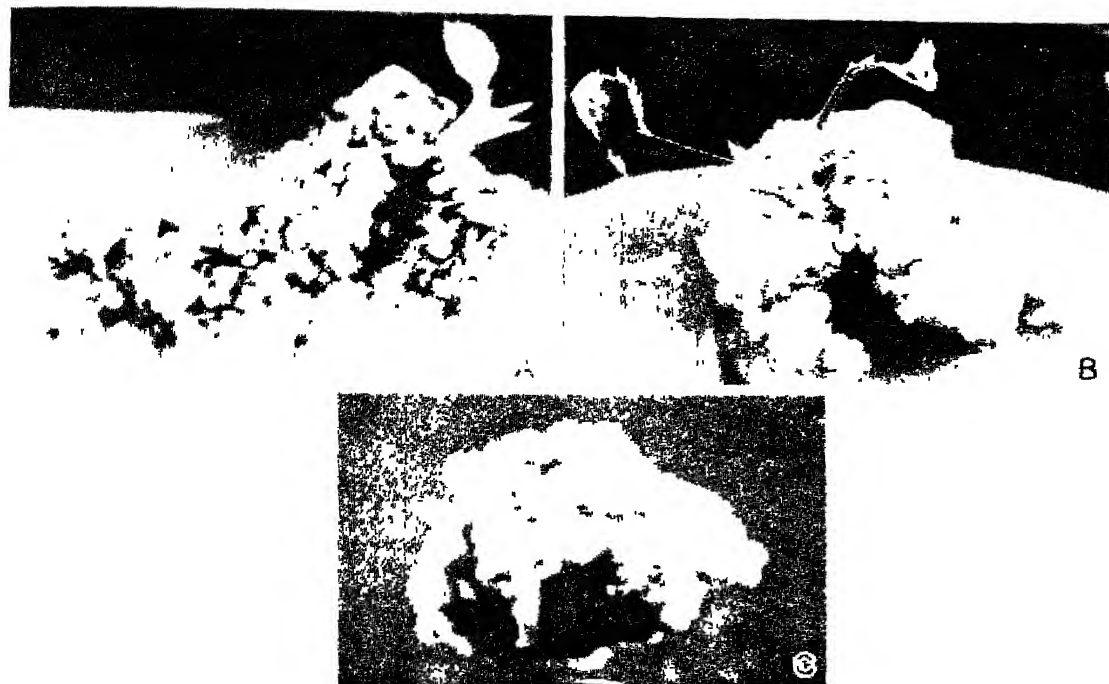


FIG. 9 A-C Apple (*Malus pumila*)

A. Shoot differentiation on stem callus on BW + NAA (4 ppm) + K (2 ppm) + CW (15%) + GA<sub>3</sub> (1ppm)

B. Shoot differentiation on shoot-tip callus on BW + NAA (2 ppm) + BAP (2ppm) + CW (15% v/v)

C Embryoids developing on cotyledon callus on BN + NAA (2 ppm) + BAP (2 ppm)

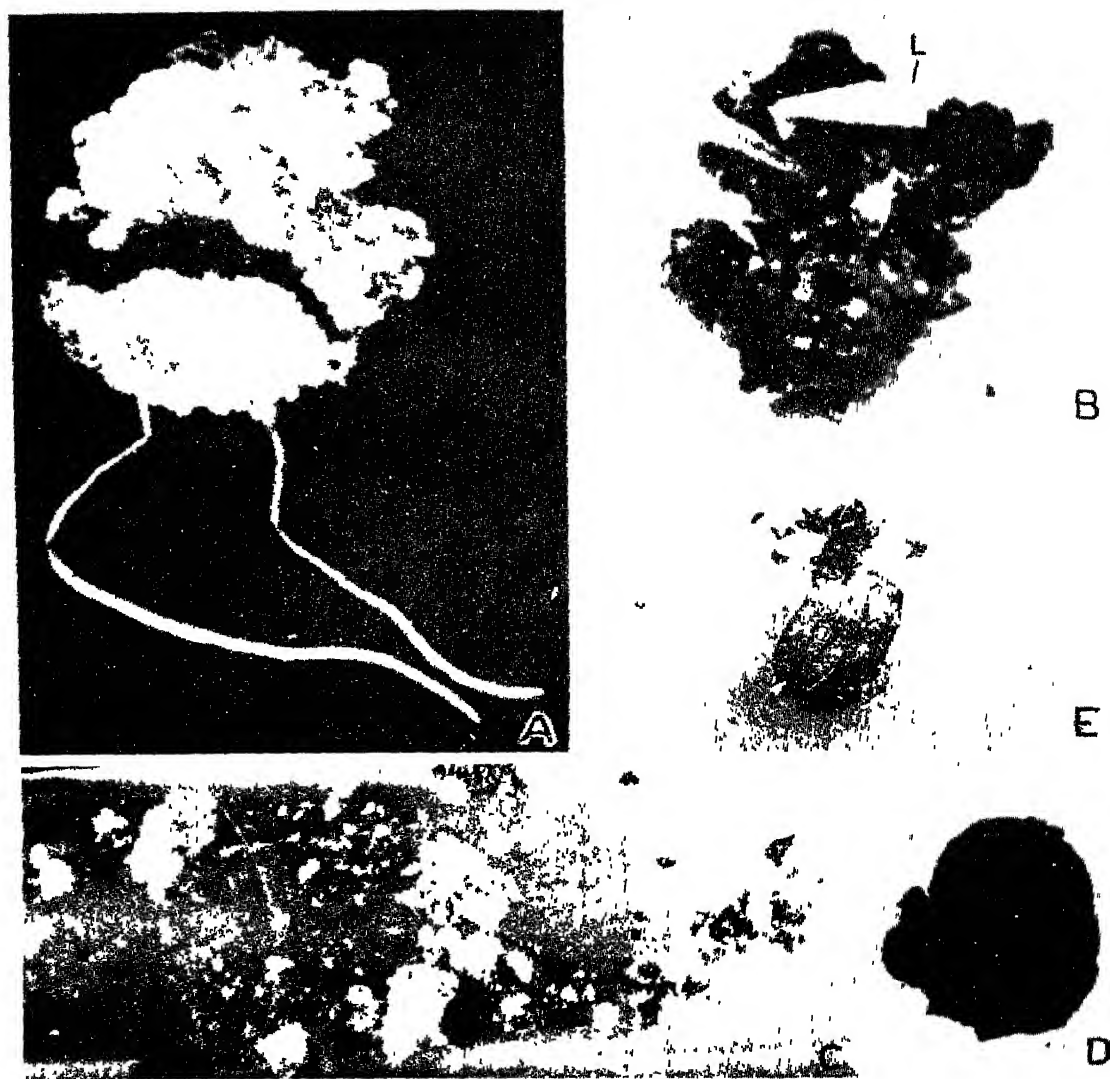


FIG. 10 A-C. Pear (*Pyrus communis* var. *communis*)

A. Root differentiating on leaf callus on BW+ Sorbitol (3%) + NAA (2.0 ppm) + K (0.2 ppm) + CH (100 ppm) + CW (10% v/v)  $\times 2$ . B. (*Pyrus communis* Var *sativa*). Isolated leaves developed from embryo callus on BW + 2.4D (2 ppm) + K (2 ppm) + CW (15%). C. Petal callus showing embryoids on BMS + NAA (2 ppm) + BAP (2 ppm). D. An embryo magnified. E. Seed germinating on BMS + K (2 ppm) showing multiple shoot formation.

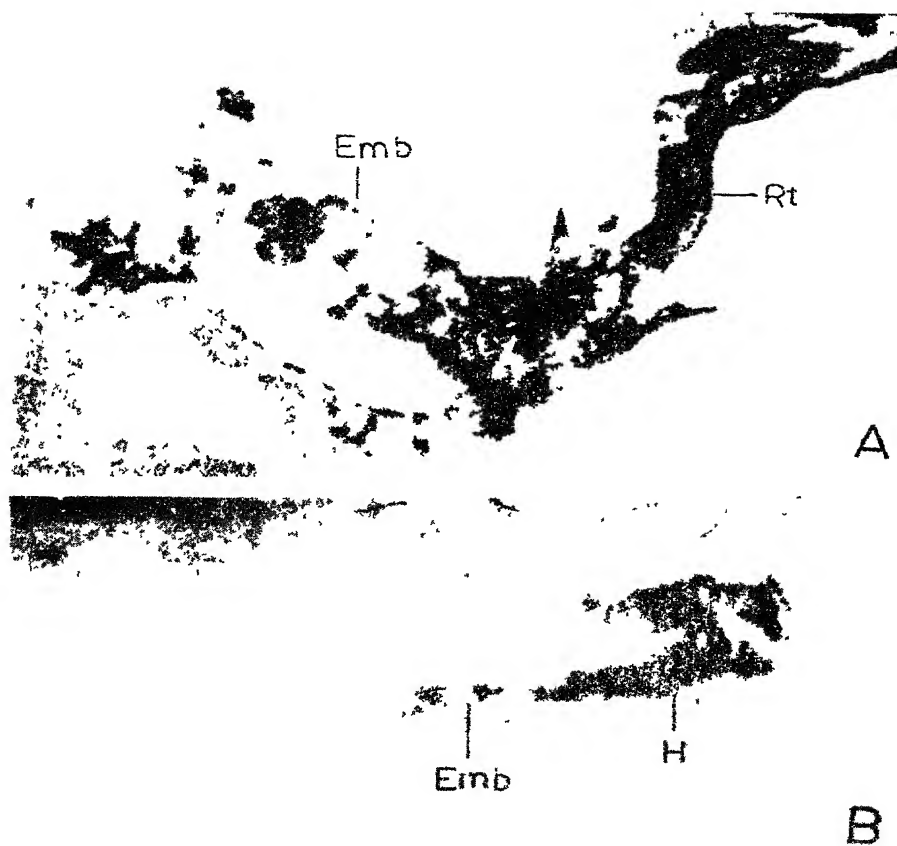


FIG 11 Chickoo (*Achras sapota*)

A. Embryoids formed directly on root explant cultured on MS + NAA (4 ppm) + K (0.4 ppm) B. Embryoids formed directly on hypocotyl cultured on BW + CW (15% v/v) + NAA (4 ppm) + K (0.4 ppm)



FIG. 12 A-F. *Pterotheca falconeri*

A. Twin root developed from shoot-apex callus or BMS + CW (20%) + IAA (2ppm) B. Isolated leaf developed from shoot-apex callus. C. Shoot developed from leaf callus D, E. Embryoids developed from ovary callus on Earle's medium + NAA (1ppm) + K (1ppm) F. Leaf differentiated from leaf callus producing secondary shoots and leaves from laminar surface

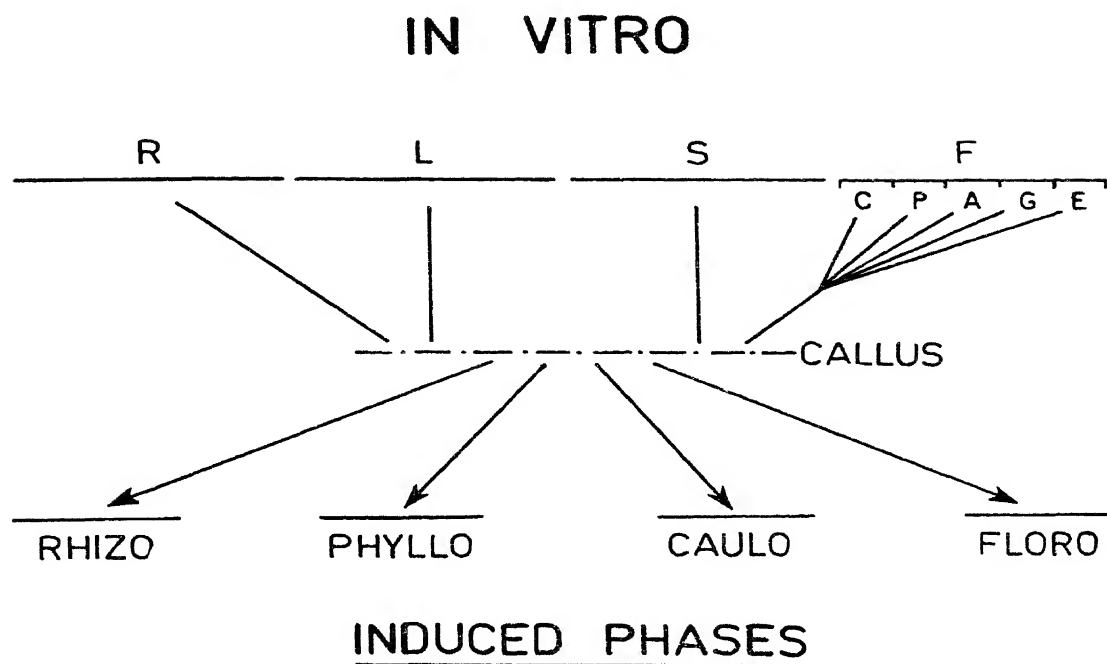


FIG. 13 Gene-block hypothesis illustrating *in vitro* development from callus cultures derived from Root, Leaf, Shoot, Flower and its parts (calyx, petal, androecium, gynoecium) and Embryo.

Diagrammatic.

R- root, L-leaf. S-stem, F. flower, C-calyx, P-petal, A. Androecium, G. gynoecium, E-embryo.

# IN VITRO

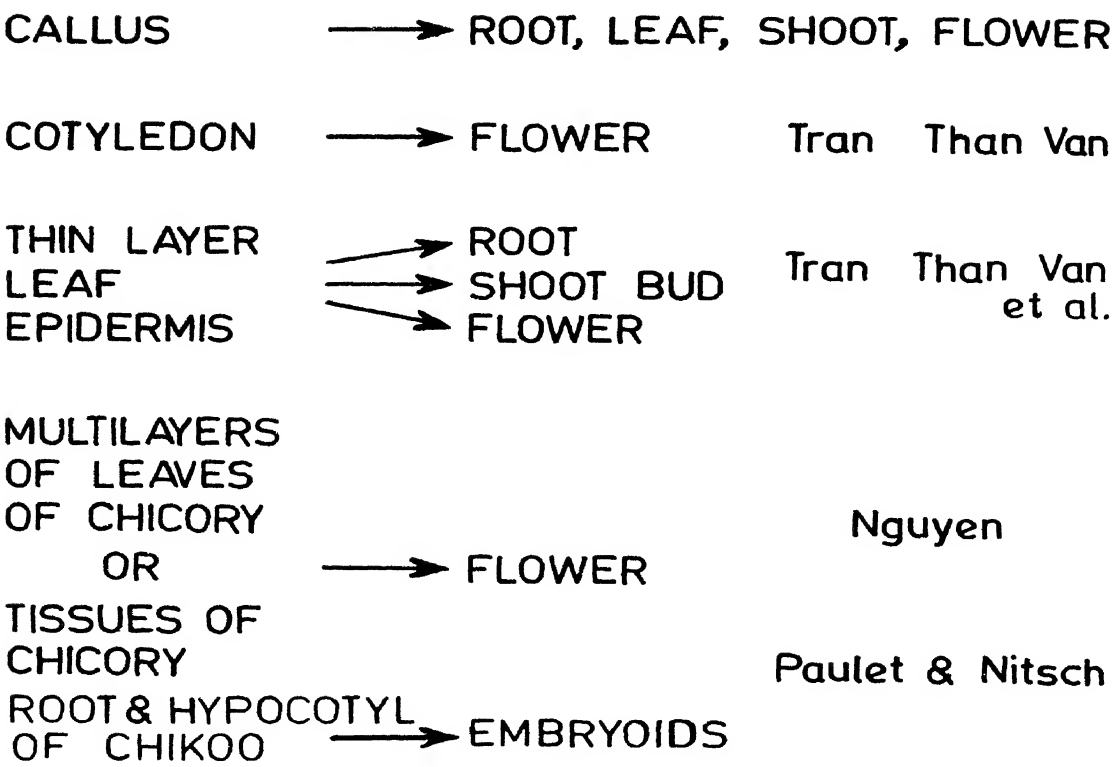


FIG. 14 *In vitro* development of plant organs from callus or various plant parts (self explanatory).



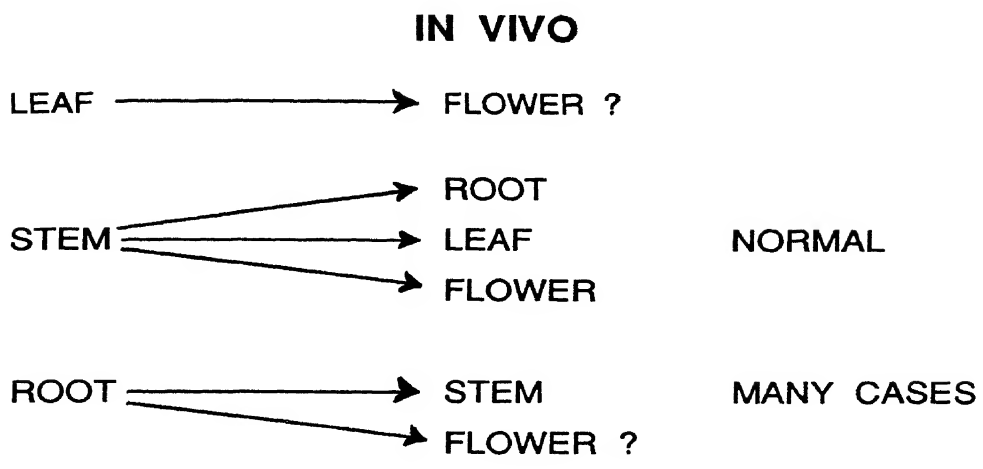


FIG 15 *In vivo* development of various plant parts

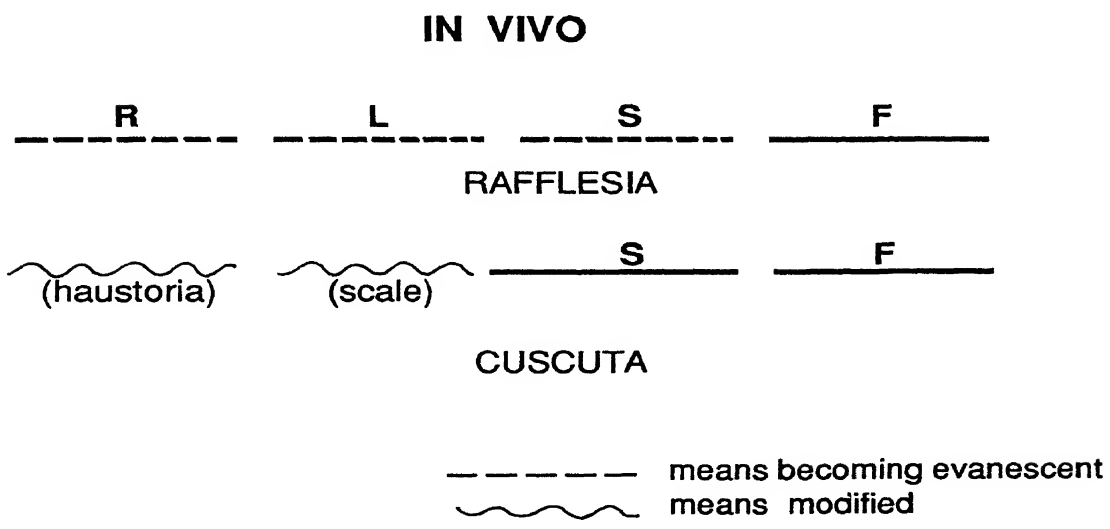


FIG 16 *Rafflesia* — *In vivo* organisation of plant parts. Diagrammatic.  
*Cuscuta* — *In vivo* organisation of plant parts. Diagrammatic

II. It has also been possible to differentiate flowers, or vegetative buds, or roots *in vitro* from thin superficial layers of floral branches of tobacco (Wis 38) by quantitative variation of carbohydrates, IBA and kinetin in the medium (Tran Than Van et al. 1974), which suggests that a particular Gene-block of the genome became activated in each case while the others remained repressed. (cf. Fig. 14).

Likewise flowers alone have been induced from multilayers of leaves or vegetative tissues of Chicory (*Chichorium intybus*) through the application of certain hormones (Paulet and Nitsch 1964).

Similarly embryoids have been obtained from superficial layers of hypocotyl or root in *Achras sapota* (Chickoo) (Fig. 11A,B). It must, however, be re-emphasised that whereas in the relatively simpler system of ferns sugar is the primary inducer for morphogenetic expressions (Mehra 1975), in the more complex system of flowering plants it is the interaction of auxins, cytokinins and sometimes gibberellins, in various concentrations and combinations that act as primary inducer to switch on specific Gene-blocks for organogenesis or embryogenesis.

#### DATA FROM NATURE (FIG. 15/2)

Roots are commonly induced from stem cuttings during establishment of plants—a phenomenon well known to horticulturists.

Shoots not uncommonly arise as regenerants from pieces of stems, leaves, or even roots in plants which are prone to vegetative methods of reproduction.

Perfect flowers complete in all respects arise in nature in *Rafflesia*, a member of the family Rafflesiaceae (parasitic angiosperms) in which root, stem and leaves, originally present, have been replaced by mycelial cord-like structures for drawing nourishment from the host. This means that transformation of the vegetative parts into 'mycelial cords' in these plants has occurred as a result of the inactivation of the root, stem and leaf Gene-blocks by the action of biochemicals present in the host root (Fig. 16).

In dodder (*Cuscuta* sps.) while the roots are replaced by haustoria and the leaves are reduced to scales, typical flowers complete in all respects arise in profusion. This, in other words, means that only the Floral Gene-block has remained unimpaired and is activated, whereas the other parts have undergone modifications (cf. Fig. 16)

Of these four Gene-blocks, the root Gene-block appears to be conservative, the stem Gene-block semi-conservative, while the leaf Gene-

block shows wide variations in nature as seen in the multifarious types of leaves in form and structure in the various angiospermic families and even within the same family in different genera and species. Floral Gene-block, on the other hand, is the most conservative and appears to be constituted of tightly cohering genes which during morphogenesis express themselves meticulously in precise sequential order, temporally and spatially, to form calyx, corolla, androecium and gynoecium. Their tight coherence is evidenced by the fact that it is impossible to induce *in vitro* any of the individual parts like sepals or petals, stamens or carpels alone bypassing the others. Likewise *in nature* a flower must appear as a whole, whether perfect or imperfect. Thus no instance is known either in *nature* or *in vitro* of the occurrence of isolated individual parts.

While the process of organogenesis according to this view from calli *in vitro* is illustrated diagrammatically in Fig. 13, the situation in regard to *Rafflesia* and *Cuscuta* as occurs in nature is depicted diagrammatically in Fig 16. In Figs. 13 and 16 the genome as a whole, for the sake of simplicity, is represented by a straight line and is divided into 4 Gene-blocks. It may be emphasised that all the genes of any one of these blocks are *functionally* integrated but may or may not necessarily lie on the same chromosome.

In retrospect and phylogenetically speaking it must be mentioned that root and leaf Gene-blocks had not come into existence in the earliest land plants—the Psilophytales. The root Gene-block is not represented even in their close living descendents, the Psilotales.

Furthermore, it needs to be reiterated once again that the Floral Gene-block is highly stable and conservative and seems to undergo only minor modifications during the evolution of species within a genus, or genera within a family. That is why the basic pattern of flower structure at generic and family level remains the same. One observes, for example, the same type of basic floral organisation barring variations in colour, size and form in all the members of any of the families, say Oleaceae, Apocynaceae, Ranunculaceae, Cactaceae, or Orchidaceae etc. The irresistible conclusion, therefore, is that the Gene-block determining the reproductive structures is not only different from that determining the vegetative characters but is rigidly stable. This stability in fact is the keystone to the classification of the flowering plants. In the event of destabilisation of a particular Floral Gene-block as a result of the

alteration of its constituent genes, the origin of higher taxa above the family level is the consequence.

## SOME BASIC PHENOMENA

### REGULATION OF GENE ACTIVITY DURING CELL DIFFERENTIATION

In the oocytes of amphibia paired lateral loops of different sizes had been observed on the so called lamp-brush chromosomes alternating with granules of different sizes. The loops are interpreted to be the sites of active RNA synthesis, while granules are inactive DNA regions of the axial strand. These granules when assuming activity open out into loops and reversely when the activity ceases they condense to form granules (Gall and Callan 1962). Izawa et al. (1963) have conclusively shown that the loops are the sites of RNA synthesis by demonstrating that they regress to granules on administration of arginine-rich histone or actinomycin D, both of which have the basic property of blocking RNA synthesis.

Occurrence of Balbiani rings or puffs at particular sites on polytene chromosomes were shown by Beermann (1963) in the cells of certain tissues in the larvae of fruit-fly and *Cheironomus*. The puffs, according to him, are opened out bands and are comparable to loops of lampbrush chromosomes and are the sites of RNA synthesis. When not in activity they regress to form bands. He also observed sequential appearance of puffs at different stages of development of the larvae. Furthermore, the cells in which puffs appeared were found to synthesise a granular secretion. Clever (1961, 63) in the same laboratory performed an elegant experiment by administering the molting hormone ecdysterone into larvae of insects and noticed the appearance of localised puffs in definite pattern and time sequence on the chromosomes. The obvious interpretation is that the hormone activated a few genes and these in turn activated some others and so on and this process occurred in regular sequential order.

That the specific RNA's are responsible for forming specific proteins has been shown by an equally elegant experiment performed by Niu et al. (1961, 62). These investigators experimented upon amphibia, mice and chicks. They extracted RNA from cells of thymus gland, liver cells, and kidney cells and this was administered to ordinary cells, and even tumor cells. The results were quite striking. They found that liver RNA induced even in the tumor cells the formation of albumin glucose-6-

phosphate and tryptophanpyrrolase which are the characteristic liver products. Kidney RNA did not induce the formation of such proteins but induced the formation of specific kidney protein L-amino-acid oxidase.

This in essence means that the genes are induced to activity in regular succession and time sequence in the Gene-blocks by the application of hormones, and these genes in turn transcribe DNA-complementary RNA which ultimately is responsible for the formation of the specific proteins.

### LAW OF PROPORTIONALITY AND THE PRINCIPLE OF SELF ASSEMBLY IN DIFFERENTIATION

This principle envisages that the macromolecules resulting from the activities of the genes constituting the specific Gene-blocs are formed in specific proportions and, furthermore, have the intrinsic property of self-assembly resulting in organisations of higher orders. It is well known that in the synthesis of insulin, two polypeptide chains with 52 amino acids in sequence in each become linked together. Many such assemblies are expected to occur in the cytoplasm and are hence under the control of cytoplasmic environment. The primary role, however, in such organisations is that of nuclear genes but not exclusively, even though they determine the basis or foundation of differentiation. The cytoplasmic environment outside the nucleus with its complexities of chloroplasts and mitochondria (each of which is endowed with its own DNA), golgi bodies, endoplasmic reticulum and ribosomes have their own vital role to play in differentiation of specific cell types and tissues in higher organisms or eukaryotes\* In this context to me the analogy of weaver's machine seems very appropriate. The basic materials like proteins (actins, tubulins and a host of others) are the outcome of gene activity, fats, lipids are the products of cytosol, sugars are synthesised by the photodynamic action of the chloroplasts, the minerals are provided by the roots, and the energy is supplied through the agency of mitochondrial activity. It is an *automatic*

---

\* Hammering's experiments in 1963 on *Acetabularia* elegantly demonstrate the cooperative role of nucleus and cytoplasm in morphogenesis in this respect.

It is also well known that the sperm in Bryophytes or Pteridophytes, constituted primarily of nuclear chromosomes, is never able to give rise to a sporophyte by itself, whereas the egg although haploid but in conjunction with the cytoplasm in which it is embedded can form embryo parthenogenetically.

*process*, and being dependent on definite parameters in respect of the quality and quantity of the specific substances formed in accordance with the genetic constitution of the organisms, there is the manifestation of differentiation with precise mathematical accuracy.

### COORDINATE INTERACTION IN THE MANIFESTATION OF STRUCTURAL CHARACTERISTICS OF ORGANS DURING MORPHOGENESIS

We could cite the example say of *Bulbophyllum reptans* or *Rhyncostylis retusa* to illustrate our point of view. Both of these species belong to the family Orchidaceae. The internal structure of the leaf in these is quite complex. There is the epidermis with paracytic type of stomata. Then there are mesophyll cells embedded amongst which are special types of water-storage cells with multispiral cellulosic thickenings. They are columnar and barrel-shaped in the former, but only barrel-shaped in the latter (Fig. 17A. B). In addition there are vascular strands with the usual organisation into xylem and phloem.

The micro-differentiation of such complex tissue of the leaf could only be interpreted on the basis of coordinate inter-action of cells in the early or primordial stages of differentiation. This could be effected by the release of the metabolites from the young primordial cells and their distribution through the plasmodesmata. These biochemicals in turn would trigger the specific gene loci and also act on the organelles within cell or cells at certain specific areas modifying them so as to form the particular type of organisation.

### CONCLUDING REMARKS

#### IN VITRO INDUCTION OF ORGANOGENESIS/EMBRYOGENESIS

Experimental evidences have shown that organogenesis/embryogenesis can be induced in every type of callus, be it from the embryonic parts of a seedling, shoot meristem, mature plant parts, or from reproductive organs like whole flowers or their individual parts. The callus has to be initiated first by culturing the excised organs on appropriate media supplemented with auxins. Such calli are amenable to differentiation depending upon the hormonal combinations to which they are subjected. In certain cases like flower parts (or stem explants prone to exudates like phlobatannins) the special substances in such calli have to be leached out before they can be made to respond to the hormones like auxins, cytokinins, gibberellins or

natural growth substances. These have to be in appropriate subtle combinations depending on individual requirements which, in turn, are governed by the genetic constitution of the tissues and their biochemical status.

As a general rule auxin/cytokinin ratio holds good for regeneration of roots, shoots, and in sporadic cases even flowers (cf. Pomegranate), but there are exceptions. These exceptions can be related to the supra-optimal synthesis of auxins (as in habituated tissues), or cytokinins, which have to be modulated by specific inhibitors so as to bring down their level to the optimum for differentiation. This is also subject to there being no chromosomal alterations of any kind.

It appears that differentiation is intrinsic to the calli. These can be geared into action to form whole plants or even isolated roots, leaves, or flowers depending on the concentration and combination of particular morphogenetic substances incorporated into the medium.

### EMBRYOGENESIS

In *typical in vitro embryogenesis* the embryoidal initial is isolated from the surrounding tissue by a mucilaginous sheath. This has been observed in a number of cases (*Punica granatum*, *Malus pumila*, *Pyrus communis*, *Achras sapota*) belonging to diverse families in the plant kingdom. This is what it should be if the embryoidal initial once activated is to be permitted to grow in its own way protected from the biochemical interference from outside. In these the whole battery of genes is initiated to activity in a precise sequential order resulting at first in a bipolar structure, the upper pole soon differentiating laterally into two first leaf primordia (Fig. 9C, 12E). Even in nature the fertilised egg is sealed from the outside influences by a similar sheath, be they ferns or flowering plants.

In contrast, in higher animals in the development of organism from fertilised egg nature operates in a grand design, far more complicated than in higher plants. The genome is parcelled out into activity quite early into adjacent tissues that initiate primordials for the programmed differentiation of the various organs.

### COORDINATION IN CELL AND TISSUE DIFFERENTIATION

Cell differentiation envisages coordinate action of all the organelles within the cell in the matter of space and time at the molecular level. The forces operative are bio-chemical, bio-physical and micro-electrochemical in the

form of ion charge. There is no organelle within a cell that has no active part to play, otherwise it would have been eliminated during the prolonged history of evolution of the organic world. The paramount role, however, is of the nuclear genes and gene-blocks, but equally important are the cell membranes, the chloroplasts and the mitochondria which are endowed with their own limited DNA.

The position and orientation of spindle in cell division is of great significance in differentiation in many cases (caulonema in mosses where the spindle is so oriented as to form oblique septa; formation of rhizoidal initial laterally in the protonemal cell in hepatics; cutting off the prothallial cells at one pole in germinating pollen grains of conifers). In unicellular organisms the position and orientation of spindle is determined by a centrosome but what governs in the higher plants is not clear.

The microdifferentiation of tissues is a highly coordinated process wherein cells constituting the tissue in the primordial stages of development interact with each other through the biochemicals released and distributed through the plasmodesmate. This interaction leads to the origin of certain special cell types like the stone cells, fibres, tannin and resin cells, cystoliths and raphides, and special type of water-storage cells with meticulously arranged and diverse types of cellulosic thickenings as seen in some orchids like *Bulbophyllum*, *Coelogyne* etc. (Fig. 17).

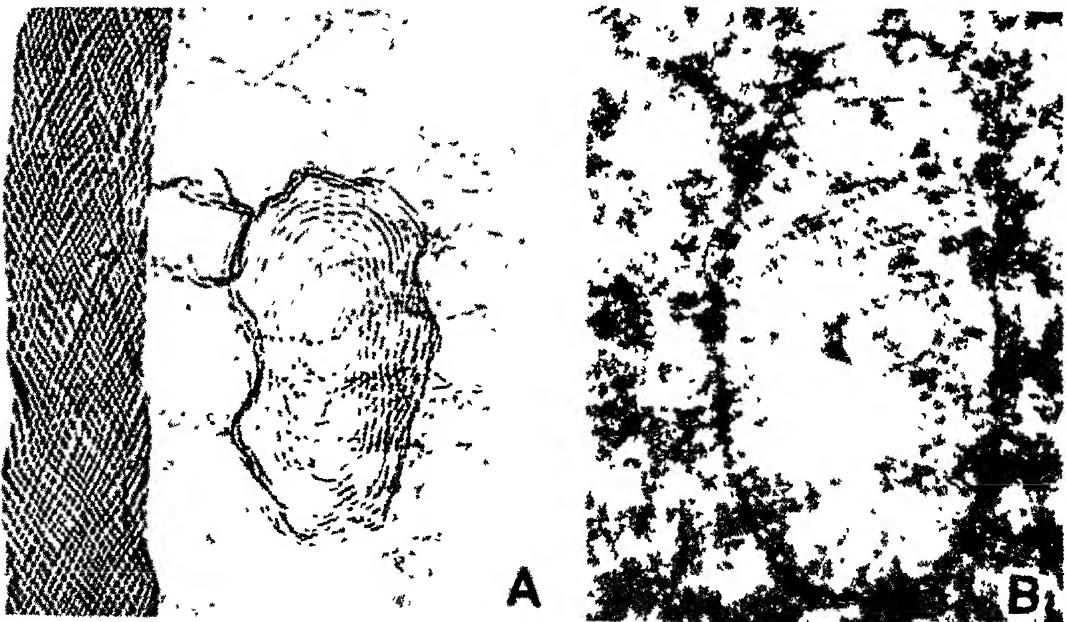


Fig 17 Water storage cells in leaf.

A. *Aerides multiflora* B. *Rhyncostylis retusa*



In short there is coordination in the functioning of organelles within a cell, coordination of cells in their organisation into tissue and the form they assume in the minutest details as for example the form and architecture, colour and texture of flowers like Panzy, Cacti, *Cassia*, *Bulbophyllum* and *Dendrobium* (orchids) which exhibit almost a mysterious symmetry in each and every flower of an individual. It is amazing how such a thing is so meticulously organised. And finally there is embryogenesis wherein the whole battery of genes are stimulated to activity in a perfectly sequential order synthesising proteins and enzymes as a result of the impact of *natural* hormones (*in vivo*) or *exogenous* ones administered to the callus tissue (*in vitro*).

### GENESIS IN DIFFERENTIATION

My personal observations on the unfolding of shoot-buds of *Mangifera indica*, *Gnephaliium litchi* and *Ficus religiosa* have led me to believe that there are seven stages of development through which the *leaves* undergo:

- |                     |   |   |
|---------------------|---|---|
| Primordial stage    | – | When there is formation of undifferentiated mass of cells.  |
| Deterministic stage | – | When there is intense metabolic activity of individual cells and correlative influences on each other leading to the pattern formation and characteristic venation. This is the most crucial stage. |
| Embryonic stage     | – | Growth and division of cells resulting in increase in size.   |
| Maturity stage      | – | When it reaches full dimension  |
| Ripening stage      | – | Intense physiological activity and synthesis of carbohydrates and other biochemical and physiological substances.   |
| Senescent stage     | – | Decline of the above processes as a prelude to decay.   |

Death and abscission

These stages are equally applicable to other organs like fruits and seeds.

The blue print of what the form of leaf, flower or fruit is going to be is accomplished at the *Deterministic stage*. I believe at this stage the entire Gene-block of the organ is in action. However, the cells embarking on different patterns of determination are not visibly different from each other in the early stages under light microscope but obviously there are switch on differences in respect of their genes. At the *Embryonic Stage*, one can make out the differences histologically. In litchi, for example, there is local eruptive activity of cells in many areas on the fruit-wall at this stage resulting ultimately in minute spines in the mature fruit wall.

### DIFFERENTIATION OF REPRODUCTIVE PARTS VERSUS VEGETATIVE PARTS

The differentiation of reproductive parts (flowers) *in vivo* is different from differentiation of vegetative parts in the sense that whereas the latter depends upon continuous elaboration of metabolites, the former comes into effect only with the advent of "adverse or unfavourable" conditions and operates primarily on the stored metabolites or metabolic reserves of plants under optimal conditions of light and temperature (cf. flowering flush in *Prunus amygdalis*, *Prunus persica*, plum and apricot, or *Bombax ceiba* and many others).

### COMMITMENT

In physiogenetic terms the action of genes/geneblocks is initiated through biochemicals or natural hormones as stated earlier. Once a tissue or organ is committed to action there is no going back. Retracing of the steps can be done only at a very early or Primordial Stage by alternation of the physical or biochemical environments. A floral-bud primordium can revert back to shoot apex if conditions necessary for its formation are altered. In ferns the leaf primordium can become transformed to a bud primordium at the formative stage before it is finally committed, but after commitment it must form the frond howsoever poorly developed it may be. The release of morphogenetic hormones at the commitment stage is such that the process becomes irreversible. In nature in *Dracaena*, *Yucca* and some other plant species, the primordium which is otherwise to form floral buds may become transformed into shoot buds—a phenomenon designated as phyllody, which, however, is genetically controlled.

A pollen cell can organise a complete embryo by switching on the *master gene* of the genome (Fig. 18) provided it has not undergone an advanced stage of its normal development along a different pathway, or in

other words has become committed. Likewise, it can also give rise to a normal embryo-sac by switching on a sub-gene of the Floral Gene-block (cf. *Hyacinthus orientalis*, Stow 1930-34; Naithani 1937; or *Ornithogalum nutans*, Geitler 1941 Vide Maheshwari, 1950). The *spermatogenous initial in the antheridia of Hepatics, Mosses, and sporogenous initial in sporangia of ferns* are likewise unchangeably committed to go along a definite pathway. This is in contrast to their wall cells which retain totipotency and are uncommitted.

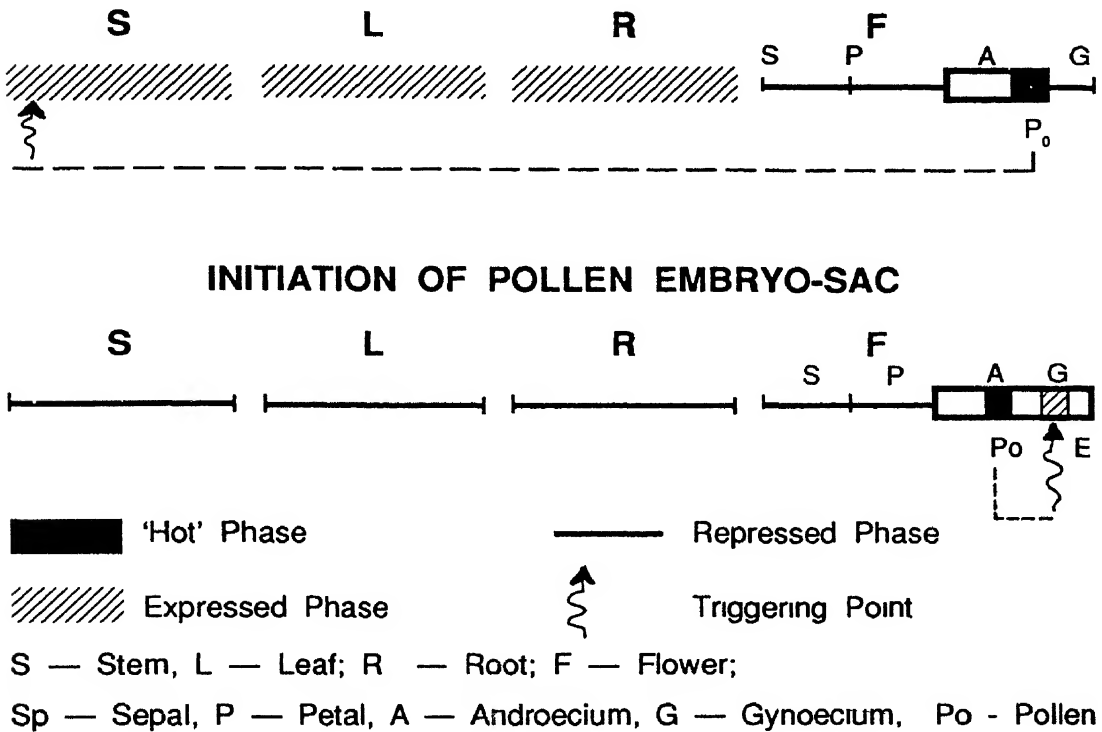


FIG 18 Development of pollen embryo and pollen embryo-sac according to Gene-block hypothesis

### SOME KNOTS OR UNRESOLVED PROBLEMS

How does an embryo-sac in a gymnosperm realise that it must form 256 free nuclei (*Ginkgo biloba*), or 512 free nuclei (*Ephedra intermedia*), or more (*Pinus* sps.), and thereafter it must stop further divisions and enter into the stage of laying down of walls to form cellular endosperm? Or how does the embryo-sac in angiosperms realise that it must enter into 8-free nucleate stage and thereafter form an egg apparatus with an egg and two synergids which must develop peculiar filiform apparatus, and three antipodals, and the two free nuclei be set aside ultimately (after fusion

with the second male nucleus) to form endospermic tissue in the *Polygonum* type of embryo-sac development? or in the *Fritillaria* type what stimulus is there which invokes the 3 nuclei at the antipodal end of the embryo-sac to fuse and thereafter undergo division ultimately to form an 8-nucleate embryo-sac with a haploid egg apparatus and a triploid andipodal complex?

Likewise an archegonial initial in all the hepatics whether Jungermanniales, Marchantiales or Sphaerocarpaceles is programmed to develop in a specific way and an antheredial initial on the *same thallus* is programmed to develop in a different way than the former. The only explanation that can be offered is that the *different subsets* of the gametophyte Geneblock have been triggered into activity in each case, but what is the underlying cause in biochemical terms?

In the case of fern gametophytes hormonal control of antheredium formation has been demonstrated by the work of Döpp (1950, 1959) on *Pteridium equilinum* and *Dryopteris filix-mas*, and Näf (1956) on *Onoclea sensibilis*. The antheredium forming factor or the antherediogen is liberated by the mature archegonia-bearing gametophytes and induces/enhances antheredia formation on young gametophytes of the same or even a number of other species belonging to many different fern families. This specific antherediogen or the *Pteridium* factor has been chemically analysed and found to be of the nature of carboxylic acid with some unsaturated linkages (Pringle, Näf and Braun, 1960; Pringle, 1961). The antheredium inducing factor in schizaceous ferns like *Anemia* and *Lygodium* is different and as shown by Schraudolf (1962, 1966), Voeller (1964) and Voeller and Heinberg (1969) can be substituted by Gibberellic acid or other Gibberellins. Nothing, however, is known about the substances inducing archegonia in fern gametophytes.

It is now clear on the basis of experimental evidences on a number of bryophyte species that gibberellin here too promotes antheredia formation. Nothing is known so far as archegonia are concerned.

If we only knew how to trigger the *master gene* of the entire genome, or the *initiator* of its sub-sets it may become in the realm of possibility to evolve a complete individual, or any of its parts. But for the expression of such genes the environment in which they have to operate must normally be favourable.

We have seen stray cases of flower development directly on the callus (Fig. 7C). Can we so manipulate that we get whole crops of flowers on the calli at will? Likewise we have seen the unusual phenomenon of origin of sporangia in sporadic cases on the gametophyte of fern *Cyclosorus* (Fig. 4B, C). Can it be possible to evoke this response in the gametophytes at will by the application of morphogenetic substances in appropriate combinations?

### RECURRENT MORPHOGENESIS (in plants) VERSUS DEFINITIVE OF PACKAGED MORPHOGENESIS (in animals)

In plants once the shoots and roots are differentiated they show *Recurrent Morphogenesis*.

Continuous growth of short-apex results in recurrent appearance of leaves at all stages of development which possess exactly the same form and configuration at maturity and arise consistently in a particular phyllotactic sequence i.e. spiral, opposite decussate, or whorled, which in other words means that the apical meristem amazingly follows a definite cyclic pattern throughout its future unfolding.

Again flowers arise in succession one after another on the same inflorescence and follow meticulously an identical pattern as regards shape, size, and configuration, as in *Cassia fistula*, *Kigelia pinnata*, or orchids and cacti. in fact in the whole of plant kingdom.

These afford unique opportunities for plant morphogenesists to understand the basic phenomenon of the working of Stem, Leaf or Floral Gene-Block or their sub-sets by the exogenous application of growth substances at different stages of their development, in vivo as well as in vitro.

But by far the most important feature in plants is the totipotency of cells of leaf, stem, flowers in angiosperms. In hepatics and mosses, the thallus, leaf, stem, stalk of the carpocephalum, involucre, cells of the antheridial wall or archegonial venter (Fig. 19), rhizoids (Fig. 20) or even stalk of sporogonium (Fig. 21) are all endowed with totipotency, exceptions being only the spermatogenous cells or sporogenous cells which are committed, as stated earlier.

In contrast animals exhibit *Definitive and packaged morphogenesis* which means *almost* simultaneous differentiation of various organs during development of the fertilised egg. High specialisation and histological

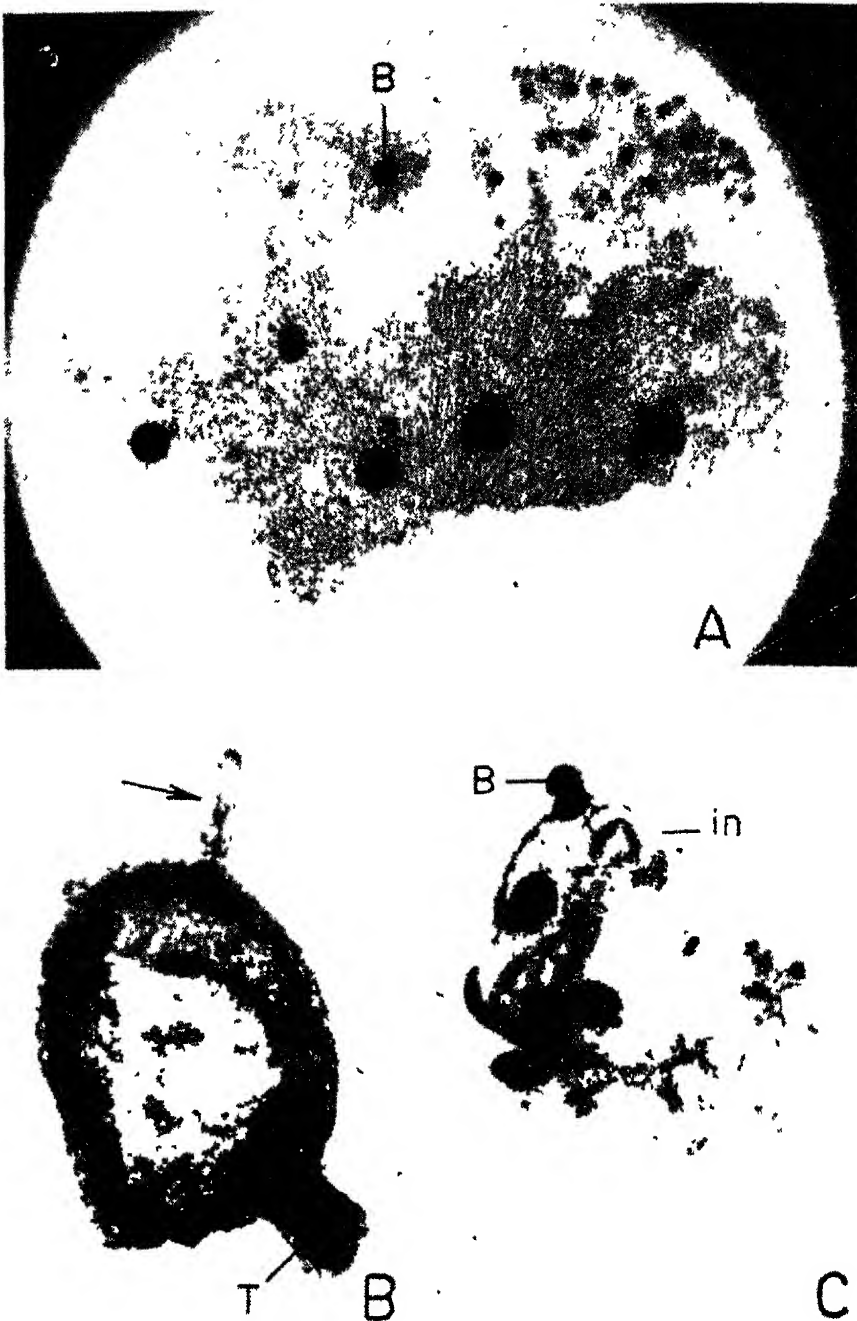


FIG 19 A. *Sewardiella tuberifera* - a wing showing early stages of regeneration of thalli in the form of buds (B). B.C. *Stephensoniella brevipedunculata* B. Calyptra regenerating a thallus (T). The neck of the archegonium at arrow. C. Involucre (In) regeneration forming buds (B).

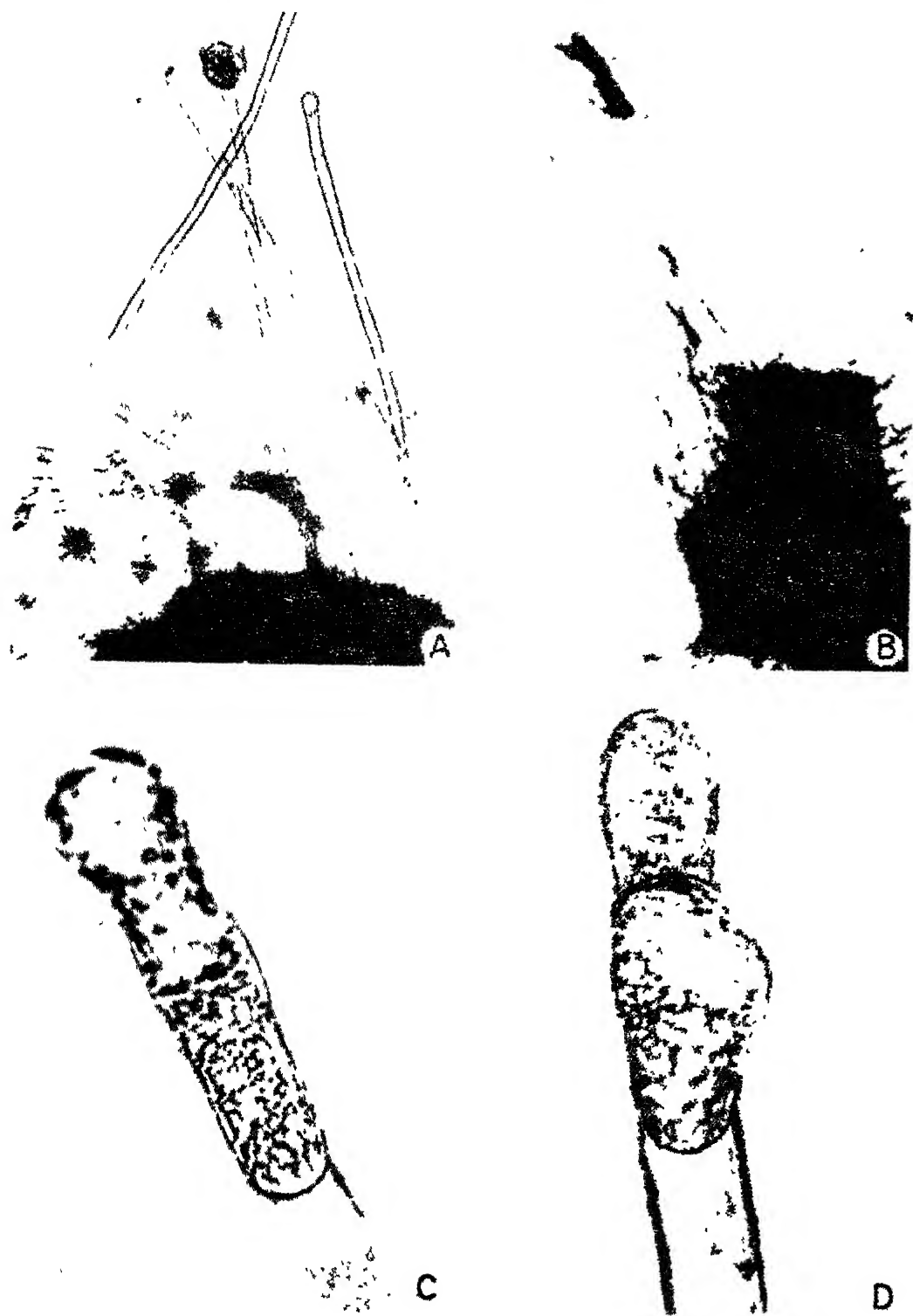


FIG 20 *Fossombronia humalayensis* A-D. Various stages in rhizoid regeneration to full plant.



FIG.21 *Athalamia pusilla* A.B. Regeneration of thalli from stalk of sporogonia



complexity of organs and tissues (especially in higher animals) results in total loss of totipotency. This is due either to specialised cytoplasmic environment (erythrocytes. nerve cells), or fundamental alterations in the nuclei (micronucleus in contrast to macronucleus in ciliates, the latter undergoing structural chromosome changes), gene amplification in chromosomes during differentiation, or casting off of genomic fragments into the cytoplasm.

## References

1. Aghion, D 1962 *C R Acad Sci Paris (Ser.D)* **255**; 993
2. Aghion-Prat, D. 1965 *Physiol Veg* (3), **229**; 303
3. Beermann, W. 1963. *Amer Zoologist*, **3**, 23.
4. Chopra, R N and Kumar, P K 1988. *Biology of Bryophytes* (p. 102) Wiley Eastern Ltd., New Delhi.
5. Clever, U 1961 *Chromosoma* **12**; 607
6. Clever, U 1963. *Develop, Biol* **6**: 73
7. Clever, U 1965 *Chromosoma* **17**:309
8. Döpp, W 1950. *Ber deutsch Bot Ges.* **63**: 139-147  
— 1959 *Ber deutsch. Bot Ges* **72**: 11-24
9. Gall, J G and Callan, H G 1962. *Proc Natn Acad, Sci* **48**: 562.
10. Hamerling, J 1953. *Intern. Rev Cytology* **2**: 475-498
11. Izawa, M, Allfery, V G and Mirsky, A E 1963. *Proc Natn. Acad Sci* **49**: 544.
12. Jacob F, and Monod, J. 1961 *J Mol Biol*, **3**: 318-356
13. Jacob F and Monod, J 1963. In *Cytodifferentiation and macromolecular synthesis* ed. M Locke New York Academic Press. 30.
14. Jacob F and Monod J 1963. Elements of regulatory circuits in bacteria. In *Biological Organisation at the cellular and supracellular level* ed. R.J.C. Harris. Academic Press, London and New York.
15. Jaidka, Kanchan and Mehra, P N *Can J Bot.* **64** 1644.
16. Maheshwari, P 1950. *An Introduction to the embryology of Angiosperms*. Mc Graw Hill Publication.
17. Mehra Asha and Mehra, P N 1974. *Bot. Gaz.* **135**: 51.
18. Mehra, P N 1975. *Research Bull. Punjab Univ. (N.S.)* **23** 221

19. Mehra, P N 1976. *J Hattori Bot. Lab.* **40**: 521
20. Mehra, P N and Cheema, G S 1980. *Phytomorphology* **30**: 336
21. Mehra, P N and Jaidka, Kanchan 1979, *Phytomorphology* **29**; 286.
22. Mehra, P N and Jaidka, Kanchan 1983, *Phytomorphology* **35**; 1.
23. Mehra, P N Mehra, Asha 1971. *Phytomorphology* **21**; 174.
24. Mehra, P N and Pahwa, M S 1976. *J Hattori Bot. Lab.* **40**; 371.
25. Mehra, P N and Palta, H K 1971 *Phytomorphology* **21**: 367.
26. Mehra, P N and Pental, Deepak, 1976. *J Hattori Bot. Lab.* **40**; 151
27. Mehra, P N and Sachdeva, Saroj. 1979. *Phytomorphology* **29**; 310.
28. Mehra, P N and Sachdeva, Saroj, 1979. *Phytomorphology* **34**; 26.
29. Mehra, P N Sulklyan 1969. *Bot J Linn Soc* **62**; 431.
30. Monod, J and Jacob, F 1961. Cold Spring Harbour Symp. Quant. Biol. **26** ; 389-401.
31. Naf. U. 1956. *Growth* **21**: 91-105  
— Naf, U and Braun. A C 1960, *Nature* **186**; 1066-1067
32. Naithani, S P 1937 *Ann Bot* **1** 369
33. Niu, M C, Cardova, C C and Niu, L C, 1961. *Proc Natn Acad Sci* **47**: 1689
34. Niu, M C, Cardova, C C and Niu, L C, and Radbill. C L 1962 *Ibid* **48**: 1964.
35. Paulet, P and Nitsch J P 1964. *Ann Physiol Veg* **6** 333.
36. Pringle, R B 1961, *Science* **133**: 284.
37. Reinert J and Bajaj V P S 1976 In Plant Cell, Tissue and organ Culture Ed. Reinert J. and Bajaj. VPS., Springer - Verlag; 251.
38. Sachdeva, Saroj and Mehra, P N 1986. *Phytomorphology*. **36**: 315.
39. Schraudolf, H 1962. *Biol. Zbl* **81**: 731-740.  
— 1966, *Pl. Cell Physiol.* **7**:277-289.
40. Sonneborn. T M 1964. *Proc Natn Acad. Sci* **51**; 915.
41. Stow, I 1930, 1934. *Cytologia* **1**: 417, **5**: 88
42. Tran Than Van. M., Chlyah, H. and Chlyah. A. 1964. In Tissue Culture and Plant Sci. ed. H.E. Street. 101
43. Voeller. B R. 1964. In *Regulateurs Natureles de la Croissance Vegetale Colloques Centres. Natl. Rech. Sci. Gif-sur-Y velte* **123**: 665-684.

- and Winberg, E.S. 1969. *In* Current Topics in Plant Sciences Ed. J. Gunckel. Academic Press.
44. Watson, J D Hopkins, N, Roberts, J W, Steitz, J A, and Weiner, A M *Molecular Biology of the Gene* Vol II.



**Chunchinkatte Melkote Srinivas Dass** (b. 20 March 1926) did Ph.D. (1955) from University of Mysore; Ph.D. (1957), University of Wisconsin, USA. He is Emeritus Professor, Karnataka State Sericulture Development Institute, Bangalore; and retired as Professor and Head, Department of Zoology, University of Delhi.

Dass demonstrated the occurrence of redundant DNA in the ciliate macronucleus and its periodic elimination during binary fission; correlation between DNA synthesis in macronuclear anlagen and degradation of macronuclear DNA in fragments in

the ex-conjugates of ciliates; paracrystalline organization of DNA in the nucleus of sperms and the transition from less basic to more basic proteins to facilitate packing density of DNA during spermiogenesis; subunit organization of ribosomes and the association of mRNA with ribosomes to form polyribosomes in rat liver cells; inter-relationship between DNA repair mechanism, macronuclear genomic content and somatic aging in ciliates; utilization of deoxynucleoside derived from macronuclear fragments in the synthesis of DNA during development of macronucleus in ciliates.

Dass is Member of Indian Society of Cell Biology. He is the recipient of Sunder Lal Hora Medal (INSA) (1990).

*Chunchinkatte Melkote Srinivas Dass was elected to the fellowship of the Academy in 1979.*

**Address:** Emeritus Professor, Karnataka State Sericulture Development Institute, Jalaghattapura Post, Bangalore Distt.; Res: 122/23 VII B Main Road, 4th Block, Jayanagar Ext, Bangalore 560 011; Tel: (080) 649269.

## NUCLEAR ORGANIZATION IN CILIATE PROTOZOA

C M S DASS

I consider it a great honour to have been awarded Dr. S L Hora Medal. When I looked into the list of previous recipients of this medal I found name of three of my illustrious teachers, and this makes me feel all the more grateful to INSA for having chosen me for this distinction.

I had the opportunity of meeting Dr. Hora way back in 1951 when he had come to Bangalore to attend a meeting in Indian Institute of Science and I was greatly struck by his towering personality. But what impressed me most was when I heard him expounding his Satpura Hypothesis to account for the occurrence of the Cyprinid fishes both in Western Ghats and in North Eastern mountain ranges. To receive this honour endowed after his name is of great significance to any biologist.

It is customary in a lecture like this to recount the work one has been engaged in during ones research career. Amongst other research interests one aspect that has fascinated me and with which I have long been associated is the Cell Biological aspects of Ciliate Protozoa especially their nuclear organisation and differentiation.

Ciliates are a fascinating group of single-celled organisms which are very different in many ways from other higher eukaryotic cells. They are simple, easy to culture and have a fairly short generation time. They are unique in that they possess two types of nuclei (macro- and micronuclei) which have different roles to play in their life cycle. I got interested in these forms when I observed in *Glaucoma pyriformis*, a very fast dividing Hymenostomatid ciliate, an unusual process of chromatin elimination from the macronucleus at the time of binary fission. While it did not occur in all cells it occurred fairly frequently in a fast growing culture. At that time the significance of this process was not clear but I will come back to this little later. Another event that fascinated me was the process of development of the macronuclear anlagen in the exconjugants of the common peritrichous ciliate *Vorticella*. It was observed that there was an apparent correlation between the development of the macronuclear anlagen, by the way of rapid synthesis of DNA, and the rate of resorption of the old macronuclear fragments in the synconjugants. It struck us that this process of development of the macronucleus from the division

products of synkaryon is of considerable significance and we formulated some questions:

1. What is the significance of the size difference between the macronucleus and micronucleus while they are derived from the same synkaryon?
2. What is the origin of this large quantity of DNA in the macronucleus?
3. Why does macronucleus not form regular chromosomes during division and why does it divide by amitosis?
4. What is the significance of the breaking up and absorption of the old macronucleus even as the new macronucleus is developing in the exconjugants?

Our work and that of many others have led to the understanding of the above questions.

Using the then recently developed Feulgen microspectrophotometry by Pollister and his school in Columbia University, Moses was the first to measure the amount of DNA in the macronucleus and micronucleus in case of *Paramecium* and established that the total amount of DNA in the vegetative macronucleus was at least 1000 times more than that in the micronucleus taking the value for that in the micronucleus as  $2c$ . This fortified the then prevalent concept that the macronucleus is highly endopolyploid and added credence to the thesis of Sonneborn that there were number of subnuclei in the macronucleus. It is now known that structurally and in function the macro- and micro nuclei differ very much. Since the micronucleus synthesizes very little or no RNA and it does not develop any nucleolar body it does not seem to have any direct involvement in vegetative functions. There are naturally occurring amiconucleate strains, one such strain of *Tetrahymena* (GL) which was originally isolated by Lwoff is still being maintained in clonal cultures in many laboratories. This strain multiplies only by binary fission. It is enigmatic, since this strain has not shown any sign of somatic senescence. Experimentally it can be shown that micronucleus can be easily destroyed either by  $\gamma$  - or UV-radiation or by chemical treatment by such chemicals as cis-Platin. They seem to be extremely sensitive to chemical treatments unlike macronucleus. In some ciliates as in *Stylonychia*, *Blepharisma* there is an attempt on the part of the organisms to produce from the macronucleus bodies resembling micronuclei - pseudomicronuclei. Why

are these bodies produced or what is the metabolic compulsion, is not clear, since they have no apparent role in vegetative functions. They also cannot participate like a normal micronuclei in producing pronuclei during conjugation. Coming back to essential differences between the two analysis of micronuclear DNA has shown that in forms like *Stylonychia* and *Oxytricha* the molecular weight of DNA is very high and does not migrate in gels showing the molecules to be very large. They are chromosome size molecules comprising of several thousand kbs. More than 70% of the sequences are unique and only about 30% show reassociation kinetics to suggest repeated sequences. There are no multiple copies of DNA. It is surmised that spacer sequences account for 95% of the sequence complexity. On the other hand, in *Oxytricha* and related hypotrichous ciliates gels of the macronuclear DNA is seen as a smear on the gel showing a continuous size ranging from 400 bp to 10,000 bp with mol wt. from  $0.27 \times 10^6$  daltons to  $6.6 \times 10^6$  daltons. Thus there is a distinct difference in the nature and organization of macro - and micronucleus in these forms. But this is not true of all ciliates. For instance, in the case of *Tetrahymena* tm of micronuclear DNA is slightly less than that of macronuclear DNA. Gorovsky has suggested 3 possibilities of organisation—macronuclear DNA (1) is in stretches smaller than DNA content in individual chromosomes, consistent with elimination of sequences from interspersed sites on most, if not all chromosomes or (2) majority of macronuclear DNA are chromosome sized, suggesting elimination of 1 or 2 chromosomes or chromosome ends or intercalated sites or (3) super-chromosome- sized molecules. Even so, there appears to be very distinct difference between macronuclear DNA and that of micronucleus. In *Tetrahymena* it has been shown that a small fraction of bases (0.6 - 0.8 mol. %) are modified to N6 methyladenine. On the other hand, it is nearly insignificant (0.1 mol. %) in micronucleus. It may even be contamination. Occurrence of methyladenine is also reported in *Paramecium*. What specific role these modified bases play in the macronuclear DNA is not clear.

We know that at the time of sexual reproduction i.e., during conjugation it is only the micronucleus that produces pronuclei. Pronuclei derived from the two participating conjugants reciprocally fuse to form zygote nucleus. While this synkaryon in each exconjugant divides and some of the division products develop to become future macronuclei, the original macronucleus in each conjugant break down and is resorbed. Very rarely some of the macronuclear fragments have shown to be capable of

producing a functional macronucleus when the macronuclear anlagen fails to develop. We have seen this happen in case of *Blepharisma* when the conjugants are treated with cis-Platin. The process of development and differentiation of the anlagen has been studied in great detail in recent years.

Our earlier photometric studies on some of the peritrichous ciliates like *Epistylis* showed that the acquisition of DNA by developing macronuclear anlagen is a gradual process linked with the possible enrichment of the nucleotide pool by the breakdown of the old macronucleus. It was several years later using  $H^3$  - thymidine labelling of the preconjugal macronucleus we were able to demonstrate in *Stylonychia* the reutilization of breakdown products, possibly upto nucleoside level, and facilitating the rapid synthesis of DNA in the macronuclear anlagen. This recycling is essential since the exconjugants do not feed and oral reorganization takes place much later. The same set of studies also showed that in these hypotrichous ciliates the differentiation of the macronuclear anlagen is complex and unique, in that most of the DNA synthesized in the anlagen initially is degraded and there is a second cycle of DNA synthesis which is linked with vegetative state. Ammermann studying the process of development of macronucleus in *Stylonychia*, by photometric measurements had also suggested the occurrence of two cycles of DNA synthesis. Furthermore, he demonstrated that the chromosomes in the macronuclear anlage at the end of first round of DNA synthesis, when they reach a c-value nearly 58 times that of micronucleus, resemble very greatly the polytene chromosomes of *Drosophila*. We still do not know how this apparent polyteny arises. Such polytene organization of chromosomes have been reported in other forms also like *Nyctotherus* but this is not true in all ciliates. Ultrastructural studies have shown that these giant chromosomes get vesiculated around each band and there is a concomitant breakdown of DNA to reach 2c level. Murti and Prescott were able to show that during this phase the chromosome size DNA molecules are chopped down to gene size molecules and processed further.

This splicing of DNA molecules is very specific and restriction endonucleases must be involved in cutting DNA molecules to appropriate sizes. This is further coupled with selective amplification of some of the gene size molecules several hundred times and simultaneous destruction of many other sequences. This amplification is associated with at least 5 rounds of DNA synthesis before the macronucleus becomes vegetatively



functional. Further analysis of these gene sized molecules in the macronucleus showed that there are about 24000 such molecules showing unique 5' terminal sequences of C<sub>4</sub>A<sub>4</sub> and G<sub>4</sub>T<sub>4</sub> sequences at 3' end as tail. These additional sequences are added later on i.e., after splicing DNA into gene size sequences. While these sequences might be associated with coding for essential functions, the DNA sequences are separate and it is estimated that there are as many as 10000 copies per macronucleus in *Stylonychia*. So it is 'endopolygenic' and the functional macronucleus is a bag of genes.

The physical organisation of the macronucleus in hypotrichs is again an enigma. Ultrastructurally they do not show any regularity in the disposition of chromatin and nearly 95% of it, at any one time, is in supercoiled, inactive 'heterochromatinized' state. Uniquely though, in the case of hypotrichous ciliates in the vegetative macronucleus at the time of DNA synthesis in preparation for binary fission the 'replication band' appears and there is a synchronous, sequential DNA replication of the DNA molecules which are randomly distributed. How this is regulated is not yet known. In fact, these are possibly the only eukaryotic cells where one can observe under phase contrast microscope the replication band and determine the duration of S-phase in the cell cycle. High voltage electron microscope observation of the replication band showed that the supercoiled DNA unwinds completely and after replication reverts to original state. Elegant autoradiographic studies of Gall showed specific basic proteins (presumably histone) reassociating with DNA in the trailing side of the replication band. Again what we have seen in hypotrichous ciliates is not true for all other ciliates. There are many ciliates where there is very little distinction between macro - and micronucleus. c-values are the same. The only distinguishing feature is the occurrence of the nucleus in the macro and its absence in the micronucleus. Again, there are several other ciliates where the macronuclear development is very different. We have found that in case of *Blepharisma* there is no DNA diminution and it does not go through DNA-poor stage. And if at all there is, it is very insignificant and not so dramatic as in hypotrichous ciliates.

Earlier I mentioned about chromatin elimination from the macronucleus in some ciliates at the time of binary fission. Now it appears that this is a process by which the errors of distribution of DNA due to unorganised movement of DNA (gene size molecules) or chromosomes when macronucleus divides by amitosis is corrected. Amitosis may have been the result of loss of 'kinetochore' which has to interact with micro

tubules for forming spindle. This might have been an adaptive step to cope with segregating thousands of copies of DNA molecules during macronuclear division. Intranuclear microtubules do appear in macronucleus but they seem to facilitate only the stretching of the macronucleus across the division plane. It is seen that in *Tetrahymena* there is 60:40 split; low amount of DNA is corrected by several rounds of DNA synthesis without reference to cell division and higher amounts are corrected by chromatin elimination. In *Stylonychia* and *Pseudourostyla* also we have found considerable variation in amounts of DNA distributed amongst daughter cells. Whether these variations are perpetuated or corrected is not clear. In *Tetrahymena* the vegetative macronucleus has about 45c amount of DNA. But in newly reorganized macronucleus it is about 64c. It takes about 50 - 100 fissions for the normalization of DNA amounts.

What can be the significance of such an enormous amplification of genes in the macronucleus. Macronucleus cannot be considered truly polyploid any more. It may be an adaptation to avoid multicellularity but to maintain the essential nucleo-cytoplasmic ratio to regulate household functions in the cell. Is this an evolutionary adaptation to amitotic segregation of DNA and to act as buffer against loss of essential genes—a form of aneuploidy? Is there a threshold level of number of copies of genes necessary for macronuclear function? Also, there is a possibility that multiplicity of copies of genes protects vegetative functions from being affected by any damage of DNA until excision—repair mechanism corrects the situation. We know from our experiments using cis-Platin there is a differential sensitivity between macro - and micronuclear DNA. We can cause extensive damage to macronuclear DNA, seen by the abnormalities that occur during binary fission and morphogenesis but gradually animals after successive generations of correction recover to normalcy. On the other hand, damage to micronuclear DNA, either artificially or during aging is not corrected. Repair mechanism does not seem to operate here. Our studies on aging in *Stylonychia* over thousands of generations has shown that while germinal aging sets in after about 100 generations as seen in the loss of ability of micronucleus to produce viable pronuclei and incapacity of macronuclear anlagen to differentiate into vegetative macronucleus to successfully restore vegetative functions, the somatic aging does not manifest itself even after 1000 generations (fission, cycles). This can be attributed to the high degree of multiplicity of

functional genes in the macronucleus and also efficient DNA damage repair mechanism in the vegetative macronucleus.

I have narrated some of the highlights of our study of ciliate nuclear organization over a period of time. I acknowledge the collaboration and hard work of my research students in these studies. I came into this field of Ciliate Cell Biology at a time when very little was known and I owe a debt of gratitude to my teacher Prof. B.R. Seshachar to have initiated me into this area.

I wish to thank you all again for this patient hearing.

## References

1. Dass, C M S Chromatin elimination in *Glaucoma pyriformis* Ehrbg *Nature* **165** (1950) 693.
2. Seshachar, B R & Dass, C M S The macronucleus of *Vorticella convallaria* Linn *J Morph* **89**, (1951) 18
3. C M S Dass & B R Seshachar. The macronucleus of *Epistylis articulata* From during conjugation - a photometric study. *Physiol Zool* **27**, (1954) 280.
4. Dass, C M S & Vimala Devi, R Action of ribonuclease on the nucleic acid system in *Spirostomum ambiguum*. *Q J M S* **103** (1962) 37
5. Sapra, G R & C M S Dass. Organization and development of the macronuclear anlage in *Stylonychia notophora* Stokes *J Cell Sci* **6** (1970) 351
6. Raminder Chadha, Sapra, G R & Dass, C M S Distribution of macronuclear DNA in *Stylonychia mytilus* - An autoradiographic study. I *J E B* **17**, (1979) 539.
7. Rajesh K Saxena, Sapra, G R & Dass, C M S Fine structure and nucleic synthesis during macronuclear development in *Blepharisma musculus* *Arch fur Protistenk* **126** (1982) 293.
8. Debjani, D Sapra, G R & Dass, C M S Action of cis-dichlorodiamine Platinum (II) on *Stylonychia mytilus* (Ciliate : Protozoa). I *J E B* **20** (1982) 4
9. Ammermann, D. Morphology and development of the macronuclei of the ciliates *Stylonychia Mytilus* and *Euplotes aediculus* *Chromosoma*, **333**, (1971) 209.
10. Nag, S F The somatic function of the macronucleus of ciliate Protozoa. *Prog. Protistol.* **1**, (1986) 219



**Sushil Kumar Mukherjee** (b. 1 January 1914) D.Sc. (1945) from University of Calcutta; He was awarded D.Sc. (h.c.) from Universities of Kalyani and Burdwan. He was Vice-Chancellor, University of Calcutta (1976-78).

Mukherjee's research has spanned varied fields: ion- exchange equilibria of clay mineral-electrolyte systems; colloid chemistry of clay minerals; membrane electrodes and ionic activity measurements; macromolecular properties of natural humus and its characterization by means of viscometric measurements; synthetic humic and fulvic acids

and their comparison with natural products; and clay-organic matter interaction in soil. His work on clay membrane electrodes and resin membrane electrodes has been widely appreciated. He has co-authored two books: 'Text Book of Soil Science', and 'Laboratory Methods in Chromosome Studies'.

Mukherjee is Fellow of Indian Chemical Society (Secretary, 1950-57), Member, Indian Society of Soil Science (sometime Secretary) and Agricultural Society of India; Hony Member, Indian Photobiology Society (President, 1977-80); President, Indian Association for the Cultivation of Science; Vice-President, Asiatic Society; President, West Bengal Academy of Science and Technology, Hony Member, International Society of Soil Science (1990) (first Indian to have received this recognition); Member, INSA Council (1988-90). He is the recipient of J.B. Chatterjee Memorial Medal (School of Tropical Medicine, Calcutta) (1980); J.C. Ghosh Medal (Indian Chemical Society) (1987); Golden Jubilee Award (Indian Society of Soil Science) (1984); Silver Jubilee Commemoration Medal (INSA) (1988).

*Sushil Kumar Mukherjee was elected to the fellowship of the Academy in 1977.*

**Address:** Flat 9, Jodhpur Court, 332 Jodhpur Park, Calcutta 700 068;  
**Tel:** (033) 473 3974

## THE SOIL NITROGEN

S K MUKHERJEE FNA

*Nitrogen is one of the essential constituents of all living things. The total N-mass is of the order of  $10^{23}$ g. Of this, 0.02 per cent, i.e.,  $2 \times 10^{19}$ g is present in the biosphere, which includes the soil. The soil N is important for plant nourishment and belongs to the biological reaction system. Cyclic processes are one of Nature's magnificent designs and three such major processes in the transformation of N may be identified. Another cycle, called the internal cycle, is specifically related to the mineralisation-immobilisation of N.*

*An inventory of gains and losses of N in soil is of great interest, but only rough estimates are possible. Biological fixation of N is one of the most fascinating ways of N-gain by soil. Microorganisms of different kinds, algae, grasses, weeds and plants have their own mechanism of N-fixation. In addition, there is chemical fixation by the soil colloids, both inorganic and organic. The soil losses are mainly effected by plant uptake, denitrification, volatilisation and leaching. There is net gain of N in soil and the form in which N accumulates in soil is mostly organic. Having combined with soil humus, N is gradually stabilised to such an extent that its mean residence time may be of the order of 200-800 years. The degree of stabilisation is such that only a very small fraction is released, but a large part of this accumulated soil N gives rise to an enormous reservoir of N in soil. It is not known how this N can be released. The chemical nature of the soil N is still unknown. But it is known to be in the form of organic compounds. As such it is speculated that some strains of microorganisms, depending on carbon for energy, may be able to unlock the huge reserve of soil N.*

Together with C, H and O, N is one of the essential constituents of all living things. Of the latter, the plants meet their nitrogen need mainly from the soil. N is a component of chlorophyll, of amino acids, proteins, essential for carbohydrate utilisation, component of enzymes, vitamins, stimulates root activity and development and supports uptake of other nutrients. The main source of soil nitrogen is the atmosphere, which contains, together with oxygen and other gases, about 79 per cent N. In this connection a rough estimate of the distribution of global N given below is of great interest. The total N mass of the globe is of the order of  $10^{23}$ g.

| N Pool      | % of total N Mass |       |
|-------------|-------------------|-------|
| Atmosphere  | ...               | 1.96  |
| Lithosphere | ...               | 97.82 |
| Biosphere   | ...               | 0.02  |

Approximately, 27,200 million tons of free  $N_2$  is present above each acre of the earth. N is present in rocks mostly as nitrides of Ti, Fe and other metals, and also as  $NH_4^+$  in silicates, and in some rocks as nitrates.

The biospheric nitrogen is present mostly in organic forms both on land, i.e., soil N, (40 per cent) and in the ocean bottom (47 per cent). 1-2 per cent mineralisation of this N supports plants and animals.

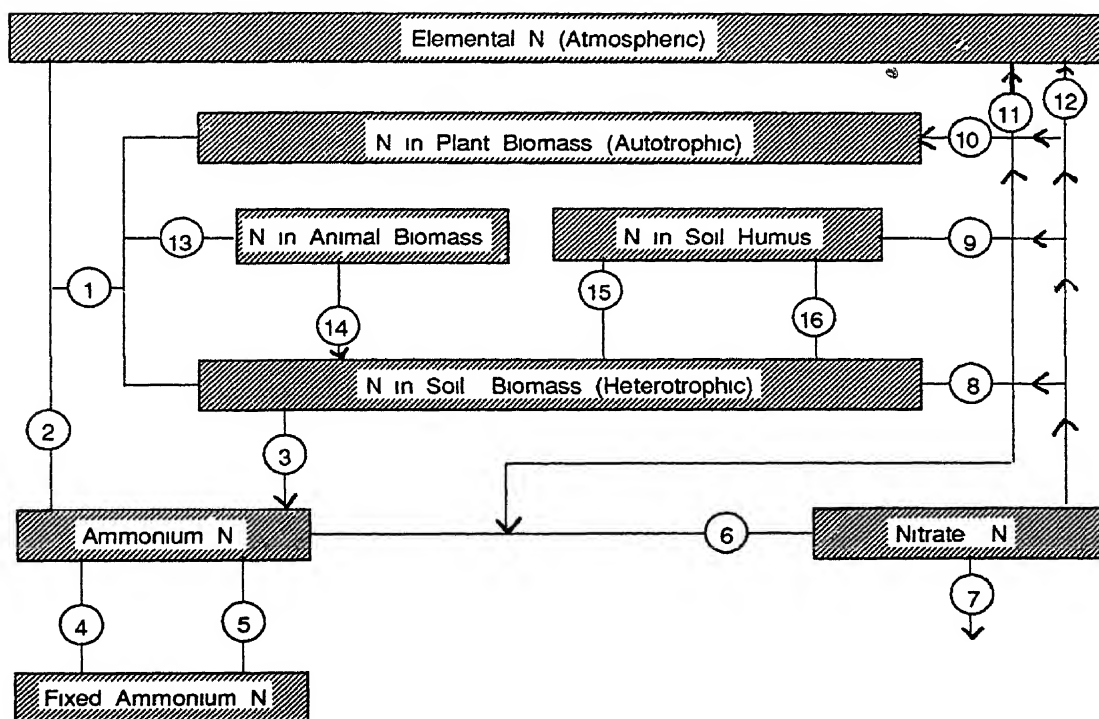
In spite of the fact that nitrogen is an inert gas, it is involved in a number of complex enzymic and chemical reactions in the soil, leading to the formation of a variety of compounds. The latter belong to amino acids, peptides, proteins, purines, pyrimidines, amides, amino sugars, nucleic acids, heterocyclic compounds derived from phenols, nitrogen-humus compound of unknown composition, and a good proportion of unidentified substances containing N. The proportions of these compounds, wherever and whenever present, vary according to the environmental conditions.

Along with P and S, N belongs predominantly to the biological reaction system, as distinct from the chemical reaction system to which other nutrients like K, Ca etc. belong. The latter are characterised by high energy of activation, obedience to law of mass action, dependence on temperature, and solution concentration etc., the last of which being altered by plant uptake. Unlike these nutrients, N requirement is not determined by plant uptake, and hence it becomes difficult to predict ability of soils to deliver N to plants. However, several transformation processes which have been identified in the soil provide some key to the information about plant need for N being satisfied. These transformation processes are fixation, nitrification, mineralisation, denitrification, immobilisation and stabilisation. These are mainly biochemical. But  $NH_4^+$  fixation by clay colloids and  $NH_3$  fixation by humus colloids are chemical in nature.

It is commonly observed that some of the biological as well as the chemical processes may occur simultaneously. In fact, at any point of time it is difficult, if not impossible, to determine the relative magnitude

of each of the processes. Even if single transformations have been extensively studied in regard to soil environment mechanisms, intermediate and end products, there are many gaps in our complete understanding of these processes. The low content of soil N was usually explained by making denitrification, estimated by difference, responsible for it. Direct measurements showed that denitrification is not so high as was made out to be. As a result, the nitrogen content of soil is not so low. But the high N-content could not also be easily explained. The story of this unexplained soil N is going to be focussed here.

Cyclic processes are said to be one of Nature's magnificent designs. The diagram given below represents the global cycle involving N. It is one of the most complex and comprehensive of the cyclic processes one finds in Nature, and show how variegated the routes of the cyclic processes are. One may identify three cyclic processes, viz., elemental, autotrophic and heterotrophic. These cyclic processes are demarcated in the diagram.



Pathways and pools in the cycling of soil nitrogen. (1) Nitrogen fixation, biochemical; (2) nitrogen fixation, industrial; (3) mineralization; (4) chemical  $\text{NH}_4^+$  fixation; (5)  $\text{NH}_4^+$  defixation; (6) nitrification; (7) leaching; (8) immobilization; (9) chemical fixation of  $\text{NH}_3$  and oxidized forms of nitrogen; (10) plant uptake; (11)  $\text{NH}_3$  evaporation; (12) denitrification; (13) animal consumption; (14) microbial consumption; (15) humus formation; (16) humus decomposition.

| Pathways included in elemental cycle | Pathways included in autotrophic cycle | Pathways included in heterotrophic cycle |
|--------------------------------------|--|--|
| Nitrogen fixation (1)                | Plant uptake (10)                      | Mineralization(3)                        |
| Animal consumption (13)              | Animal consumption (13)                | Nitrification (6)                        |
| Microbial consumption (14)           | Microbial consumption (14)             | Immobilization (8)                       |
| Mineralization (3)                   | Mineralization (3)                     |  |
| Nitrification (6)                    | Nitrification (6)                      |  |
| Denitrification (12)                 |  |  |

One can distinguish between the primary pool of N, the atmosphere, and the secondary pool, namely, animal biomass of dead organisms, humus from plant biomass, and  $\text{NH}_4^+$ -N. The last one is derived from the primary pool by fixation through predominantly biological and, to a small extent, chemical processes.  $\text{NH}_4^+$ -N may again (i) be immobilised by heterotrophic microflora; (ii) be nitrified to be taken up by autotrophic plants; and (iii) react with soil colloids to get fixed and form the pool of unavailable N.

One can identify another cycle, called the internal cycle of N, in which mineralisation-immobilisation processes predominate. This can be enlarged by including another cycle, to account for loss of N by uptake, denitrification, leaching and humus formation.

The primary pool yields N, as already stated, to soil mainly by the process of biological fixation, which is generally recognised as nonsymbiotic and symbiotic fixation. Amongst the nonsymbiotic fixers are a host of microflora, such as BGA, photosynthetic bacteria, rhodospirillum, aerobic bacteria : azotobacter, Beijerinckia, Derxia and anaerobic bacteria : clostridium, actinomycete : Frankia in association with nonleguminous plants. The claims made by workers about the amount of N fixed by free-living bacteria vary considerably. Perhaps lower figures are nearer the actual values. But in addition to the above, about 100 organisms are recognised to be N-fixers, whose contribution to the N pool of soil may be appreciable. The rhizosphere of many grass sods has been found to be rich in N fixed from the atmosphere. Gains in soil N through BNF may be seen from the following data.

| Conditions   | N gains in Kg/ha/yr |       |
|--|---------------------|-------|
| Soil amended with crop residues                                | ...                 | 15-78 |
| Field plots under sod like crop                                | ...                 | 14-56 |
| Lysimetric studies   | ...                 | 25-67 |
| Stands of pines sp. or other monoculture nonnodulated trees... |                     | 36-67 |



Amongst the higher plants the Leguminosae family fixes in varying amounts atmospheric nitrogen in root nodules in symbiosis with the appropriate rhizobia. There are about 10-12 thousand plant species belonging to this family, out of which only 1200 species have been studied for nodulation. About 90 per cent of them produce nodules and nearly 100 species are utilised for food production. All of the thousands of plants, besides those cultivated, contribute substantially to N-content of the soil.

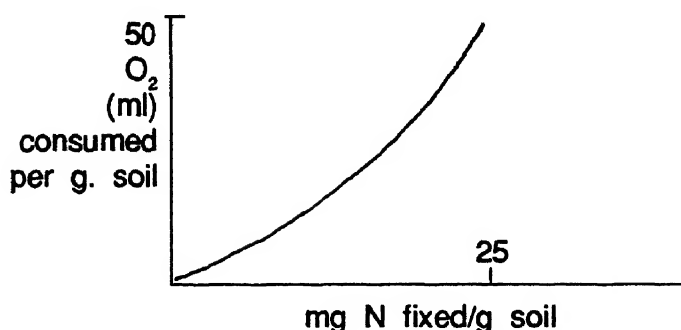
In the area of symbiotic BNF there is ample scope of genetic manipulation. For instance, development of N-fixing bacteria capable of living on the roots of such cereals as wheat; development of rhizobium strains that are insensitive to soil  $\text{NH}_4^+$  and  $\text{NO}_3^-$  concentrations; development of legumes having photosynthetic capability so that energy for N-fixation is provided to the bacteria in the nodules.

$^{14}\text{C}$  and  $^{15}\text{N}$  studies have shown that 11-17 per cent of biologically fixed N is taken up by plants, which have a preference for  $\text{NO}_3^-$  and to a small extent for  $\text{NH}_4^+$  by some plants. The rest of N remains in soil as organic N. This may slowly become available, but not immediately. In course of time, the fixed N is converted into more stable forms, rendering availability of N very low.  $^{15}\text{N}$  study has revealed that 10-40 per cent of applied N remains in soil. Plants remove 4-6 per cent of  $^{15}\text{N}$  applied. After the second growing season 24 per cent  $^{15}\text{N}$  remains behind, only 1.5 per cent of which becomes available to the plant, the remaining portion being in equilibrium with native soil N present as humus N. The mean residence time for the first year of added N is about 5 years. With the progress of humification of carbonaceous matter the mean residence time (MRT) may increase to 25 years. The MRT of N in native soil humus may vary from 200-800 years, and a small amount may remain unavailable for centuries.

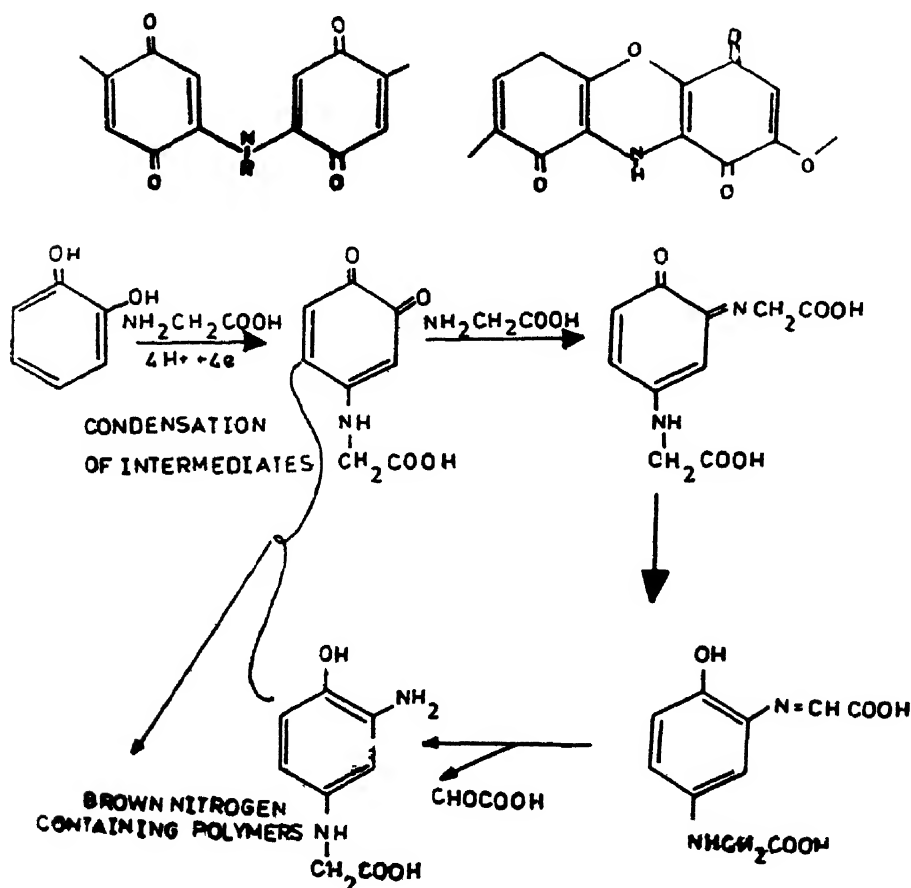
All soil colloidal fractions which predominate in clay minerals, besides oxides of iron and aluminum and humic substances, are capable of reacting with  $\text{NH}_4^+$  ion derived from the biologically fixed N or N-fertilisers. With the inorganic colloids the binding may be so strong that both leaching loss and biological availability are reduced considerably. Continuous extraction with  $\text{K}^+$  may defix a major quantity of  $\text{NH}_4^+$  and give an idea of the measure of fixation.

Organic colloids, particularly humus, react with  $\text{NH}_4^+$  in an unknown manner. At high pH  $\text{NH}_4^+$  gives rise to  $\text{NH}_3$  which reacts with phenolic or their oxidation products leading to stabilisation of N. Fixation of

$\text{NH}_3$  by organic matter is associated with oxygen uptake as shown by the figure.



Reaction with phenols and formation of heterocyclic compounds, and nitrogen bridges are the other possible chemical reactions. Radio C dating has shown that soil fulvic and humic acids may have residence times varying from a few hundreds to a few thousands of years. In a dynamic soil system, at any point of time, three fractions of organic constituents could be identified:



(1) biomass turnover at least once every year; (2) microbial metabolites and cell wall constituents that become stabilised in soil and possess a half-life of 5-25 years; and (3) the resistant fraction which in grassland soils are composed of humic substances ranging in age from 250 upto 2500 years.

In general, the availability of chemically fixed N is relatively low, of the order of 5 per cent.

The ranges in and magnitude of  $\text{NH}_4^+$  fixed by surface soils are shown by the following data collected from different countries of the world.

---

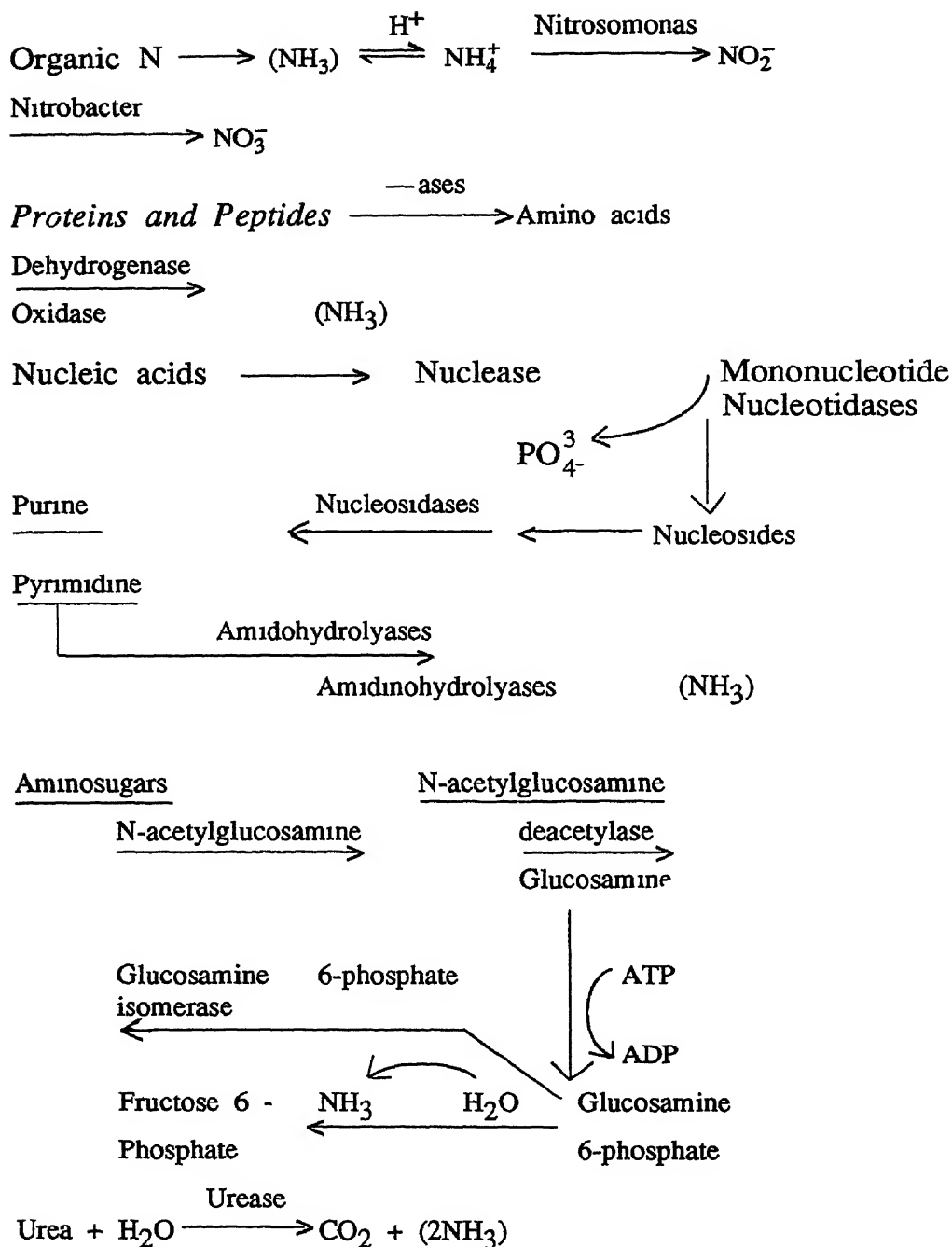
|           |     | mg/g*     |
|-----------|-----|-----------|
| Australia | ... | 41 - 1076 |
| Canada    | ... | 158 - 330 |
| Alberta   | ... | 110 - 370 |
| England   | ... | 52 - 252  |
| Nigeria   | ... | 32 - 220  |
| Russia    | ... | 14 - 490  |
| Sweden    | ... | 10 - 17   |
| Taiwan    | ... | 140 - 170 |
| USA       | ... | 7 - 270   |
| Hawaii    | ... | 0 - 585   |

---

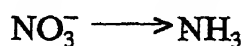
\*1mg/g = 2.24kg/ha

Measurements on rhizosphere soil reveal that the volume of such soil may contain as high as 1,700kg/ha N.

The key feature of the internal cycle of N in soil is immobilisation (organic N)  $\longleftrightarrow$  mineralisation ( $\text{NH}_3$ — $\text{NO}_3$ ). Organic soil N, as already mentioned, is composed of a variety of substances, such as plant and animal residues, microbial biomass, including partially stabilised dead organic matter, and stable humus fraction. Once this last stage is attained, mineralisation stops but is compensated for by the formation of new humus material. Many of the organic compounds in soil are broken down by the appropriate enzyme systems into simpler compounds and finally to ammonia. This ammonia is then acted upon by specific microorganisms leading finally to  $\text{NO}_3^-$ . Some of these reactions are described as follows:



Some and often all of  $\text{NH}_4^+$  and  $\text{NO}_3^-$  are simultaneously consumed by heterotrophic microflora and converted into microbial tissue. In this way the inorganic nitrogen forms are immobilized. The mechanism is the reverse of mineralisation.



$\text{NH}_3 + \text{C} \rightarrow$  amino acids, proteins, purine, pyrimidine, aminosugars, nucleic acids.

Since mineralisation ( $M$ ) and immobilisation ( $I$ ) occur simultaneously one talks of net ( $M$ ) and net ( $I$ ). Energy for either of the processes is derived from carbonaceous sources, and hence a qualitative relation is observed between C/N ratio and net gain or loss of  $\text{NH}_3$  and  $\text{NO}_3^-$ . Thus, it may be generally stated that

$\text{C/N} < 20$  Net gain of  $\text{NH}_3$ ,  $\text{NO}_3^- (M)$

$\text{C/N} = 30$  No net gain or loss, hence equilibrium

$\text{C/N} > 30$  Net loss of  $\text{NH}_3$ ,  $\text{NO}_3^- (I)$

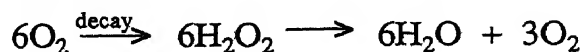
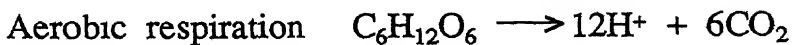
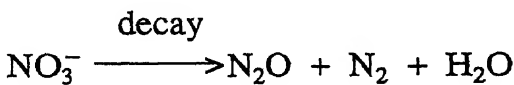
From the following data on C/N of some organic materials, one can easily judge whether they favour mineralisation or immobilisation.

| Material                         |     |     | C/N ratio |
|----------------------------------|-----|-----|-----------|
| Microbial tissue                 | ... | ..  | 6-12      |
| Sewage sludge                    | ... | ... | 5-14      |
| Soil humus                       | ... | ... | 10-12     |
| Animal manures                   | ... | ..  | 13-25     |
| Legume residues and green manure | ... | ... | 13-25     |
| Cereal residue and straw         | ... | ... | 60-80     |
| Forest wastes                    | ... | ... | 150-500   |

Soil gains in N take diverse ways. So also do soil losses in N. The following processes are generally ascribed to cause N loss in soil.

Anaerobic respiration corresponding to aerobic respiration is in fact bacterial denitrification.

Anaerobic respiration  $\text{C}_6\text{H}_{12}\text{O}_6 \longrightarrow 12\text{H}^+ + x \text{CO}_2$  (+ acids & alcohols)



Most of N of the atmosphere have passed at least once through this denitrification cycle.

For bacterial denitrification, organic compounds, reduced S compounds or  $H_2$  act as electron donors; anaerobic or restricted  $O_2$  condition; supply of  $NO_2^-$  or  $NO_3^-$  to serve as terminal electron acceptors. Not all the heterotroph bacteria belonging to the Agrobact, Bacillus and Pseudomonas are capable of producing  $N_2$  but some produce  $N_2O$ . The environmental condition should ensure poor drainage, temperature  $25^\circ C$  or higher, near neutral pH, good supply of decomposable organic matter.

In addition to denitrification, atmospheric N is enriched with  $NH_3$  derived from volatilisation from various sources.

| Sources          |     | $NH_3$ loss ( $10^9$ kg/yr) |
|------------------|-----|-----------------------------|
| Wild animals     | ... | 2-6                         |
| Domestic animals | ... | 20-35                       |
| Fuel combustion  | ... | 4-12                        |

Percentage losses of  $NH_3$  from added fertilisers are of the following order :

| Fertiliser      |     | % loss |
|-----------------|-----|--------|
| Urea            | ... | 3.5-40 |
| $(NH_4)_2 SO_4$ | ... | 3-50   |
| $NH_4 NO_3$     | ... | 17     |

The variations are due to pH, texture and organic matter content of soil.

After having looked into the ways by which the soil gains and loses N, it may be possible to make an inventory of N. But since the transformations and processes involving N are of a dynamic nature, it may not be possible to get the real picture of soil N unless measurements are made over long periods of time. On the other hand, it seems advisable to separate the soil organic components by chemical means, and find out if the N content of each fraction so extracted and chemical nature identified add up to the total. This simple procedure has revealed fractions that are not biodegradable to any appreciable extent and one fraction at least that does not only biodegrade but defies chemical identification.

This fraction has been labelled as the carrier of "Unknown" nitrogen. The results of these kind of experiments are summarised below.

*Soil N Distribution by Acid Hydrolysis*

| From                            | Method  | % soil N<br>(usual range) |
|---------------------------------|---|---------------------------|
| Acid insoluble N                | Soil N remaining after 6N HCl hydrolysis<br>(Usually by difference)   | 20-35                     |
| NH <sub>3</sub> -N              | MgO distillation of hydrolysate   | 20-35                     |
| Amino acid N                    | Determined preferably by the ninhydrin -<br>NH <sub>3</sub> method  | 30-45                     |
| Aminosugar-N                    | Steam distillation with phosphate-borate<br>[buffer (pH 11.2) corrected for NH <sub>3</sub> -N]<br>(Hexosamine N)                                     | 5-10                      |
| Hydrolysable Unknown N<br>(HUN) | Not accounted for as NH <sub>3</sub> , amino acid or<br>amino sugars [Part occurs as non $\alpha$ -amino N,<br>arginine, histidine, lysine & proline] | 10-20                     |

Acid insoluble N may comprise N as bridge linking quinone groups as described earlier. Amino acids attached to aromatic rings may not also be acid hydrolysable. Similar data on distribution of soil N collected from widely different climatic zones are as follows :

|                    | Arctic<br>(6)* | Cool<br>Temperate (82) | Subtropical<br>(6) | Tropical<br>(10) |
|--------------------|----------------|------------------------|--------------------|------------------|
| % Total Soil N     | 0.02 - 0.16    | 0.02 - 1.06            | 0.03 - 0.30        | 0.24 - 1.61      |
| Acid insoluble N** | 13.9 $\pm$ 6.6 | 13.5 $\pm$ 6.4         | 15.8 $\pm$ 4.9     | 11.1 $\pm$ 3.8   |
| NH <sub>3</sub> -N | 32.0 $\pm$ 8.0 | 27.5 $\pm$ 12.9        | 18.0 $\pm$ 4.0     | 24.0 $\pm$ 4.5   |
| Amino acid—N       | 33.1 $\pm$ 9.3 | 35.9 $\pm$ 11.5        | 41.7 $\pm$ 6.8     | 40.7 $\pm$ 8.0   |
| Aminosugar—N       | 4.5 $\pm$ 1.7  | 5.3 $\pm$ 2.1          | 7.4 $\pm$ 2.1      | 6.7 $\pm$ 1.2    |
| HUN                | 16.5           | 17.8                   | 17.1               | 17.6             |

\* No. of samples examined

\*\* This and the rest as % of total N

Cropping has a tendency to increase HUN, apparently at the expense of amino acids as the following data show.

| HUN        |      | Acid<br>insoluble N<br>(%) | NH <sub>3</sub> -N<br>% | Amino<br>acid-N<br>(%) | Amino<br>sugar-N<br>(%) | HUN<br>(%) |
|------------|------|----------------------------|-------------------------|------------------------|-------------------------|------------|
| A. Virgin  | (10) | 25.4                       | 22.2                    | 26.5                   | 4.9                     | 21.0       |
| Cultivated | (10) | 24.0                       | 24.7                    | 23.4                   | 5.4                     | 22.5       |
| B. Virgin  | (4)  | 20.8                       | 19.8                    | 44.3                   | 7.3                     | 7.8        |
| Cultivated | (4)  | 19.3                       | 24.5                    | 35.8                   | 7.0                     | 13.4       |

The percentage distribution of N is somewhat altered in the presence of added fertiliser N, as shown by the following data.

|                              | Fertiliser derived<br>organic N (%) | Native humus N(%) |
|------------------------------|-------------------------------------|-------------------|
| Acid insoluble N             | 10.3                                | 21.7              |
| NH <sub>3</sub> -N (organic) | 10.6                                | 18.1              |
| Amino acid—N                 | 59.0                                | 36.0              |
| Aminosugar N                 | 9.9                                 | 8.0               |
| HUN                          | 10.9                                | 16.2              |

The acid insoluble N, NH<sub>3</sub>-N and HUN are comparatively large in native humus. It is likely that amino acid N which shows an appreciably high value in fertiliser derived organic N is converted into acid insoluble N, HUN and NH<sub>3</sub>-N.

Both acid insoluble N and HUN are reservoirs of organic N of unidentified nature accumulated in soil over many many years. The acid insoluble part has been further extracted with 0.1 N NaOH solution and acidified to pH 2, and again extracted with 0.1 N NaOH and dialysed to give two fractions which contain about 15% or so of N. These fractions contain the locked-up N, called the 'unknown' N, but should be as thoroughly purified as possible in order to identify their N-compounds. It has been estimated that the total quantity of nitrogen so locked-up in the soil as unknown N is sufficient to supply the N requirement for crop production for 25-40 years at the present rate of demand. If this untapped source is made to yield N at an appropriate rate, the problem of N-supply will cease to exist. With the rising level of understanding of genetic engineering is it not pertinent to ask the question : Should it not be possible to produce by genetic manipulation strains of microorganisms which would unlock the N and release it from its hidden recess in the soil ? We will perhaps never unravel the exact nitrogen containing entities present in soil and the nature of N binding in them. In that case empirical trial and error methods may be the easiest and only way.

Our finger has been pointed to the mighty reservoir of N in the soil. Let us all search for the key that is going to unlock that reservoir.





**Madhu Sudan Kanungo** (b. 1 April 1927) Ph.D. (1958) from University of Illinois, Urbana, USA. He is Emeritus Professor (for life), CSIR, Emeritus Scientist (1989- ) Banaras Hindu University, Varanasi; Director, Institute of Life Sciences, Bhubaneswar (1989- ); was formerly Visiting Professor, West Virginia University, USA (1978).

Kanungo has done in-depth research to elucidate the biochemical and molecular mechanism of aging in animals. He was the first to show experimentally that (i) isoenzymes of enzymes change, but (ii) the primary structure of enzymes does not change

during aging. So he proposed through a model that the structures of genes that code for proteins do not change but their expression change after adulthood. As a result, the levels of proteins (enzymes) alter and there is a deterioration of function leading to aging. His recent studies on several genes support this hypothesis. He has authored two books: 'Biochemistry of Aging' (Academic Press, 1980), and 'Genes and Aging' (Cambridge University Press, 1993).

Kanungo is Fellow of Indian Academy of Sciences and National Academy of Sciences (India); President, Indian Association of Gerontology (1981-88). He is the recipient of Shanti Swarup Bhatnagar Prize (1971); UGC National Lecturer (1982-83); FICCI Award (1989), Jawaharlal Nehru Fellow (1987- 89); UGC National Fellow (1976-78); Third Age Award (International Congress of Gerontology, Mexico) (1989), Golden Jubilee Commemoration Medal (INSA) (1992).

*Madhu Sudan Kanungo was elected to the fellowship of the Academy in 1975.*

**Address:** Emeritus Professor, Zoology Department, Banaras Hindu University, Varanasi 221 005; and Director, Institute of Life Sciences, 301 Saheed Nagar, Bhubaneswar 751 007; Tel: (0542) 310291/354 (Varanasi), (0674) 57795 (Bhubaneswar); Res: 9 GF, Old Medical Enclave, BHU, Varanasi, Tel: (0542) 311984.

## **MOLECULAR BIOLOGY OF AGING**

**M S KANUNGO**

### **INTRODUCTION**

I wish to express my gratitude to the President and the Council of the Academy for the honour they have done me by conferring the Golden Jubilee Commemoration medal. I have devoted the best part of my research career towards the understanding of the molecular mechanism of aging with the objective that once it is known it may be possible to extend the period of youthful life. Moreover, aging is an intellectually challenging biological problem that awaits a break-through.

Aging as a phenomenon in the life spans of organisms has intrigued mankind from time immemorial. Why and how is it that having attained a vigorous adulthood, all functions should deteriorate? The duration of this phase which is referred to as aging or senescence, varies with species. It can be as short as a few days as in the female octopus which lays eggs only once, broods them, reduces its food intake and dies soon after young hatch. This is a sort of "sudden death" phenomenon that occurs soon after one-time reproduction, and the duration of aging is too brief to be perceptible in this species.

In most species, however, such a phenomenon is not seen. Mice and rats give birth to a large number of young ones, take care of them during their weaning period, and are ready to breed again soon after. In higher mammals such as humans and elephants, only a few young ones are produced during the entire life span with long gaps, sufficient to take care of the young ones during the crucial early developmental period. These species have long life spans, and live long after reproduction has ceased. For example, female rats stop reproducing after about 1.5 years, but they may live thereafter for another 1.5 years. Human females do not reproduce after 45 years, but they may live up to 100 years.

Life span is a continuum. At one end is growth and development, and at the other end is deterioration of functions or aging. In-between is the reproductive phase or adult-hood. The time of onset, duration and rate of aging are dependent on the vigour and vitality of the reproductive phase, and those of the reproductive phase are dependent on the vigour

and vitality of the growth phase. The three phases are thus inter-related. Hence aging or senescence should not be considered as an independent phase of the life span. Information on the mechanism of developmental and adult phases may help in our understanding of the mechanism of aging.

During growth, both the number and sizes of cells increase. This is followed by their specialisation or differentiation to perform specific functions. The sizes of the organs as also the size of the organism increase. Reproductive ability is attained after a certain period of growth. The reproductive phase is characterised by the appearance of factors necessary for reproduction such as sex hormones. The rate of reproduction is high initially and then gradually declines. Faster the reproduction rate, shorter is the generation time of a species. Or larger the number of off-springs produced, shorter is the life span of the species. For example, mice and rats reproduce much faster than larger mammals like elephants, horses and human beings, and they have shorter life spans. It appears as if reproduction depletes the organism of some essential substances which are not replenished as fast as they are lost. Whether this loss is the cause of aging and shortening of the life span is not known. The duration of the reproductive phase is more or less defined, particularly in females. In female rats reproductive ability is attained at about 10 weeks of age and ends at about 70 weeks. In the human female, it starts at about 12 years and ends at about 45 years.

Aging is a characteristic of all multicellular organisms. It is marked by a gradual decline in the functional ability of the organism which becomes perceptible towards the latter part of the reproductive phase. The reproductive phase smoothly merges into the phase of aging when the organism does not reproduce and the activities of all organs decline, some declining earlier than others. Several changes occur at the molecular level resulting in decreases in the functions of organs and of the whole organism. The duration of senescence is not well defined as it is not known at what stage of the life span deterioration of functions of various organs begins. If cessation of reproductive ability is used as a criterion, then it begins at 70 weeks in female rats and at age 45 years in human females. However, in human males and females, several functions like the muscular activity and breathing capacity begin to decline even from age 30. The senescence phase is of little consequence for the perpetuation and evolution of the species as the organism does not reproduce.

## THE IMPORTANCE OF STUDYING THE PROBLEM OF AGING

In the early part of this century, infectious diseases such as respiratory and intestinal diseases that were caused by micro-organisms, accounted for most of the deaths in the human population. With the advent of antibiotics against these micro-organisms, and other advances in medical sciences, these diseases have been more or less controlled. The great benefit from these advancements is the increase in the average or mean life span of the Indian population from age 30 at the time of independence to about 58 years now, and from age 40 in 1900 to age 76 in the U.S.A now. The control of these diseases has not, however, made human beings escape from aging and death. The main killers during the last two decades have been cardio-vascular diseases, cancer and cerebrovascular diseases in the order mentioned.

The rapid strides made in genetic engineering, molecular biology and medical sciences have given hope to man to reach the maximum life span of 100 years. Nearly 15% of the population in developed countries is now over age 60, while their average life span is about 76. By the year 2000, the percentage of the population over 60 years of age shall be about 20. In the developing countries including India, due to improper health care, infectious diseases, malnutrition and undernutrition only about 6% of the population is over 60. This percentage is continuing to increase due to improvement in medicare. One of the effects of birth control measures and attempts to attain zero population growth rate, particularly in developed nations, is that the population is getting older; and a smaller proportion of younger people is supporting an increasingly larger proportion of older people. Older people are susceptible to diseases, are physically less capable of carrying out their work and need financial, social and medical support. The rapid increase in the number of old people in every nation and the problems associated with it have generated a sense of urgency and interest in all nations to tackle various aspects of the problem of old age.

Aging is a universal phenomenon, and the discovery of the cause of aging shall be a fundamental finding. The aim of the researchers is to "add life into years, not years into life". The questions that the researchers are trying to answer are: Why do all organisms show deterioration of functions after attaining reproductive maturity? Why do all members of a species have a more or less fixed life span? Why does a rat live for only three years, an elephant 70 years, and a man 100 years? At what age after attaining maturity does the process of aging begin? Is there a trigger, a

switch, which sets in motion the process of deterioration? If so, how is it switched on ? Or, are there multiple switches? Is this process programmed ?

Answers to the above questions may help in designing experiments to postpone or defer the onset of the aging process, and thus control aging. This would prolong the present active and youthful period from 20 to 40 years to say 20 to 60 years. This would greatly increase the active period in humans, besides giving them the satisfaction of being youthful longer. The objective is to ensure better health for longer period, and not prolongation of the years lived. Prolongation of the youthful period is also expected to postpone or defer the onset of 'old age' diseases like cardiovascular and cerebro-vascular diseases, cancer, arthritis, etc. which set in largely between age 40 and 50 when the persons are at the peak of their careers.

Observations such as: (a) fixed life spans of all individuals of a species; (b) more or less a similar pattern in deterioration of various functions in all organisms after a short period of reproduction; (c) long life span of progeny of long-lived parents and short life span of progeny of short-lived parents, and (d) similar life span of identical twins, indicate that the cause of aging lies at the level of genes. However, factors like nutrition and heredity, and stresses of various kinds such as temperature, pollution, radiation and social and psychological stresses influence the rate of aging. This accounts for the variability in the rate of aging and life spans seen among individuals of a species. Elucidation of the changes that occur in genes after the reproduction phase or adulthood may throw light on the basic cause of aging.

### ENZYME CHANGES DURING AGING

Enzymes carry out various reactions in the cells of the body. They are coded by genes. Also, many enzymes occur as dimers or tetramers, and exist in multiple molecular forms or isoenzymes. An important example of such an enzyme is lactate dehydrogenase (LDH). It is made up of two types of protein chains, H and M, each being coded by a separate gene. It exists as tetramers and hence has five isoenzymic forms, H<sub>4</sub>, H<sub>3</sub>M<sub>1</sub>, H<sub>2</sub>M<sub>2</sub>, H<sub>1</sub>M<sub>3</sub> and M<sub>4</sub>. LDH is essential for production of ATP in the cell in anaerobic condition. Studies in our laboratory have shown that M<sub>4</sub>-LDH which is more efficient for energy production in anaerobic condition decreases with increasing age in the heart, brain and skeletal muscle

(Singh & Kanungo, 1968). This makes these tissues more dependent on oxygen and, therefore, more vulnerable to functional failure when oxygen supply is cut off due to blockage of blood supply by a clot in the blood vessel. This is one of the reasons for the higher frequency of heart attack and stroke in old age. Another important inference from this study is that the decrease in M4-LDH in old age is possibly due to the decrease in the activity or expression of the gene coding for the M-subunit of LDH.

Another interesting finding is that several essential enzymes such as acetylcholinesterase and cholineacetyltransferase which are necessary for conduction of nerve impulse at the synapse get induced by steroid hormones,  $17\beta$ -estradiol and testosterone. The levels of these enzymes decrease with age in the brain of the rat. However, administration of  $17\beta$ -estradiol to old rats restores them to adult levels (James & Kanungo, 1978). The steroid hormones, on entering the cell, bind to promoter regions of specific genes and regulate their expressions. So it is possible to maintain the desired levels of enzymes in an old animal for its optimal performance.

Proteins have been purified from young and old animals and their primary structures examined by peptide mapping and immunological studies. It has been found that their structures in both young and old animals are the same (Kanungo & Gandhi, 1972). Since the primary structures of enzymes are coded by respective genes, and the former do not change with age, the primary structures of their genes or the DNA also do not change. The alterations in the levels of enzymes that are seen with increasing age are, therefore, due to changes in the regulation of the expression of their genes.

### CHANGES IN GENE STRUCTURE

A gene which codes for a protein has a coding sequence which is split into exons (sequences that are expressed) and introns (sequences that are not expressed). The coding sequence is preceded by a promoter region which has several short sequences (*cis*-acting elements) that are involved in the regulation of the expression of the gene. These sequences, on binding to proteins (*trans*-acting factors) cause conformational changes in the promoter leading to specific interactions with the transcription factors (proteins) that bind to the gene to carry out transcription or production of messenger RNA (mRNA). This interaction determines the rate and level of mRNA production, and thus regulates gene activity.

Besides the above, it has also been shown that methylation of cytosine (C) in 5' -CCGG-3' sequences located either in the promoter or in exons and introns, and regions of the gene that are hypersensitive to the endonuclease, DNase I (DH-sites), are involved in gene activity. Generally, if -CCGG- is methylated at internal cytosine (-C<sup>m</sup>CGG-), a protein binds to this sequence and inhibits the activity of the gene. Usually the occurrence of DH-sites in a gene signifies that it is active, as inactive genes have no DH-sites. Thus, using the above criteria it is possible to find out what changes occur in a gene that make it inactive/active during aging.

Another factor that influences the activity of a gene is the degree of its complexity with the chromosomal proteins. Genes (DNA) are located in chromosomes inside the nucleus of a cell, and are complexed with proteins called histones which have positively charged amino acids. Because DNA has negative charges due to the presence of phosphate groups, histones bind to the DNA to form DNA-histone complexes to form bead-like nucleosomes along the long DNA molecule. This 'beads on a string' like structure is called chromatin which is the basic or primary structure of chromosomes. Another group of proteins, non-histone chromosomal (NHC) proteins, is also present in the chromosomes. In order for a gene to be transcribed into an mRNA it is necessary for histones to dissociate partially from the DNA so that RNA polymerase is able to glide along the DNA molecule reading its sequence and transcribing the mRNA. If the chromatin becomes compact or condensed, the speed of transcription is retarded. Such a situation may arise in old age, especially in cells like the neurons of the brain, cardiac and skeletal muscle cells, in which DNA synthesis stops very early during development and the cells stop dividing.

The possibility of the chromatin of the brain becoming increasingly compact with increasing age was studied by digesting it by DNase I that cuts chromatin DNA at 10 base pair (bp) intervals and its multiples. It was found that less of 10 and 20 bp DNA fragments are produced from the chromatin of the brain of old rats (Fig 1; Chaturvedi & Kanungo, 1985a). This clearly shows that the chromatin becomes more compact with increasing age.

Another method of testing whether the chromatin becomes compact or not is to carry out replication or DNA synthesis or nick-translation using labelled nucleotides. When DNA in the chromatin was nicked or cleaved by restriction endonucleases, EcoRI or Msp I or Hpa II, and then

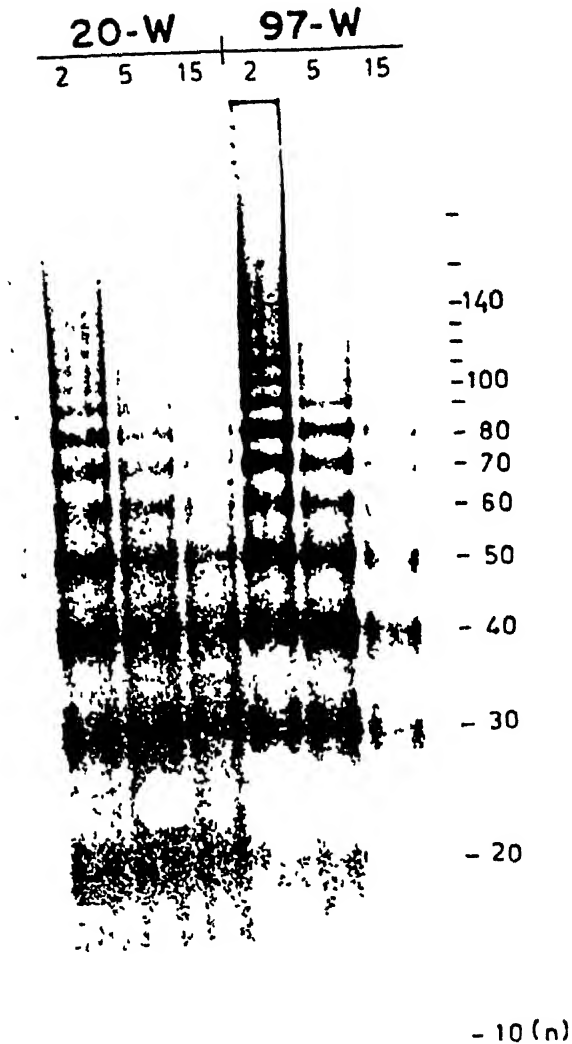


FIG 1 DNA fragments resolved on 12% denaturing polyacrylamide gel after digestion of nuclei of cerebral hemisphere of 20- and 97-week old male rats by DNase I for 2, 5 and 15 min. n-nucleotides: w-week (Chaturvedi & Kanungo, 1985a).

DNA synthesis was measured, the incorporation of the label was far less in the old (Chaturvedi & Kanungo, 1985b). These studies taken together clearly established that one of the major causes of the decrease in the expression of genes in non-dividing cells during aging is the increasing compaction of the chromatin.

### CHANGES IN GENE EXPRESSION

Studies on the albumin gene of the liver of the rat have revealed that the coding region of the gene of the young rat is more sensitive to DNase I than that of the old. Furthermore, the coding region of the gene has a -CCGG- sequence, and its digestion by the restriction enzymes, Msp I and



Hpa II, reveals that this sequence in the old is not cleaved by Hpa II. Msp I cleaves -CCGG- and -C<sup>m</sup>CGG- sequences, but Hpa II cleaves only -CCGG- sequence. This clearly shows that the -CCGG- sequence gets methylated at internal cytosine in old age (Singh, Singh & Kanungo, 1990).

The degree of transcription of the gene was examined by purifying RNA from the liver of young and old rats, and hybridizing it with <sup>32</sup>P-labelled albumin cDNA by dot blot (Fig.2). A significant decrease in hybridization was observed in the old as compared to that of the young. These studies show that transcription of the albumin gene decreases as a function of age, and methylation of its -CCGG- sequences and its insensitivity to DNase I contribute to its lower expression. Thus, structural changes in the albumin gene cause decrease in its expression as the animal ages.

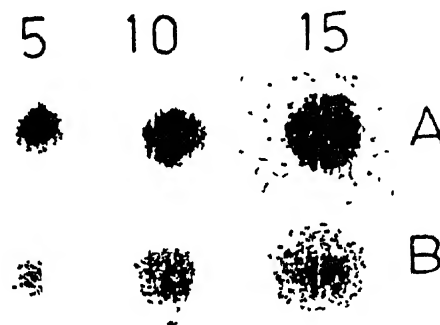


FIG 2 Dot-blot hybridization of total RNA with <sup>32</sup>P-labelled albumin cDNA. Level of albumin mRNA in the liver of young (A) and old (B) rats (Singh, Singh & Kanungo, 1990).

Another gene of much importance is the fibronectin (FNT) gene which is expressed in the liver. Its product is FNT protein which is secreted to the blood for transportation to other organs. FNT protein is made up of two subunits, each of 220 to 240 kilodaltons. They are joined by -S-S- bonds at the C-terminal ends to form the dimer. FNTs are involved in differentiation, migration and adhesion of cells, and also wound healing, hemostasis and tumour metastasis (Hynes, 1990). A single FNT gene is present per haploid genome in rodents and humans. The gene is located in chromosome 2 in the rat, spans ~70 kilobasepairs (kb) and has 50 exons. Subunit variants of FNTs are generated by alternative splicing of primary transcripts. The promoter region of the gene has several *cis*-acting elements including TATA, CCAAT and GGCGGG, and responsive elements for cyclic AMP (CRE), glucocorticoid (GRE) and

heat shock (HSRE) that are involved in its regulation under appropriate conditions (Fig.3).

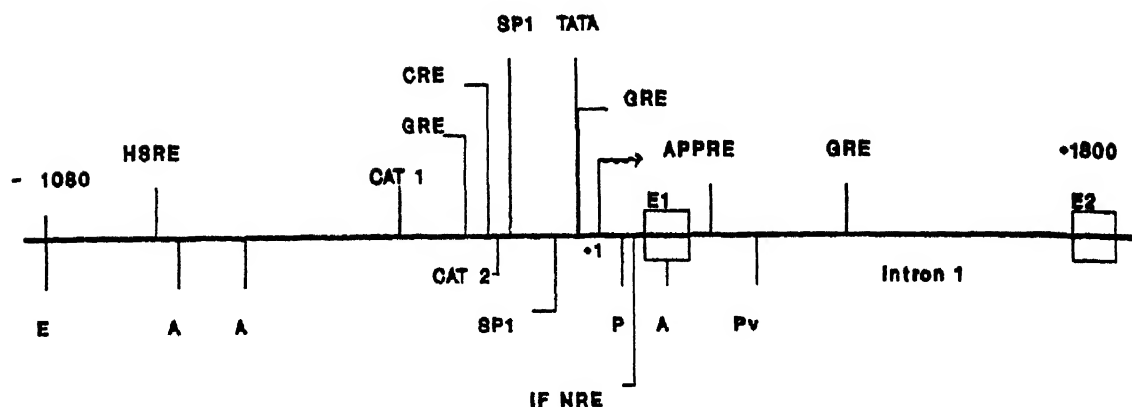


FIG 3 Map of the fibronectin gene showing *cis*-acting elements, restriction sites, first two exons and first intron. E - Eco RI; P -Pst I; Pv -Pvu II; A - Ava I; E -Exon, +1 - Transcription start site, CRE -cAMP response element; GRE - glucocorticoid response element; CAT -CCAAT; HSRE - heat shock response element; APPRE - acute phase protein response element; IFNRE - interferon response element; Sp 1 - Sp 1 binding site.

Studies on the expression of the FNT gene of the rat as a function of age has been carried out by Kanungo and his associates. They have shown that the level of FNT protein in the plasma of a 125 week old rat is significantly lower than that of a 20 week old rat. This is due to a lower rate of transcription of the gene. When nuclear run-on transcription was carried out with nuclei from the liver of young and old rats and  $^{32}\text{P}$ -labelled UTP and other NTPs, and the transcripts were hybridized to the FNT complementary DNA (cDNA) probe by slot-blot, the signal for the old rat was far lower than that for the young rat. Furthermore, digestion of nuclei by DNase I followed by restriction digestion of DNA fragments and Southern hybridization with a 1.2 kb probe encompassing the promoter region of the gene show that there are three DH-sites, one overlapping the CRE, the second in the TATA region and the third in the first intron (Singh & Kanungo, 1991). The DH-sites in the promoter are more susceptible to DNase I in the young than in the old. FNT is not synthesised in the brain in which it is neither transcribed nor its promoter has any DH-sites. There is no difference in the methylation status of -CCGG-sequences of the promoter region of the gene between the liver of young

and old rats as seen by Msp I/ Hpa II digestion of nuclei followed by Southern hybridization with a probe for the promoter region. Thus, the lower transcription of the FNT gene in old rats is due to greater compaction of the chromatin encompassing the gene and its lower sensitivity to DNase I. Methylation of its -CCGG- sequences is not a factor for its lower transcription.

Transcription of the FNT gene was studied by purifying RNA from the liver of immature, adult and old rats and hybridizing it to its labelled cDNA by slot-blot hybridization (Fig. 4; Singh & Kanungo, 1993). A gradual decrease in the level of mRNA was seen with increasing age. Also, when dexamethasone was administered to the rat and then the level of mRNA was measured, an increase in its level occurred in all ages, but it was the lowest in the old. The promoter of the gene has a GRE site which is involved in its induction by dexamethasone. The hormone binds to a receptor protein in the cytoplasm to form a H-R complex which then binds to the GRE of the FNT gene located in the chromosome in the nucleus to cause its induction. It is likely that the level of the receptor decreases with age, and hence the induction also decreases.

Our studies on the role of the cAMP responsive element (CRE) on the regulation of the FNT gene as a function of age have thrown much light on the molecular mechanism involved in the expression of the gene. The CRE has the sequence -TGACGTCA-. A synthetic 25-mer oligonucleotide containing the CRE as shown below was labelled with  $^{32}\text{P}$ .

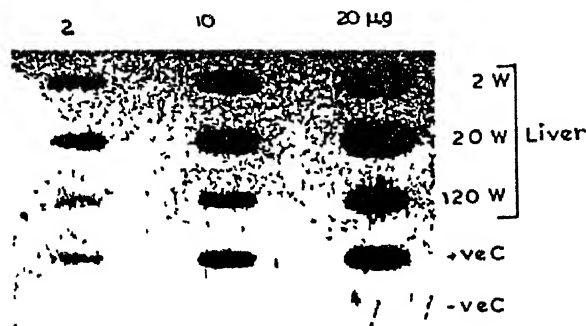


FIG 4 Slot-blot hybridization of total RNA with  $^{32}\text{P}$ -labelled FNT cDNA. 2, 10 and 20  $\mu\text{g}$  of total RNA from liver was slot-blotted on nytran membrane and hybridized to cDNA of fibronectin. +ve control: 0.1, 0.5 and 1.0 pg cold probe. -ve control: yeast tRNA (Singh & Kanungo, 1993).



It was incubated with increasing amounts of nuclear extracts of immature, adult and old rats to allow binding of CRE binding proteins (CREB) to the CRE. This was followed by gel mobility shift assay. Also, the same amount of 25-mer DNA was labelled and incubated with a given amount of nuclear extract of each age. It was then titrated with increasing amount of cold 25-mer DNA and resolved by gel mobility shift assay (Fig. 5; Singh & Kanungo, 1993). This technique is based on the principle that

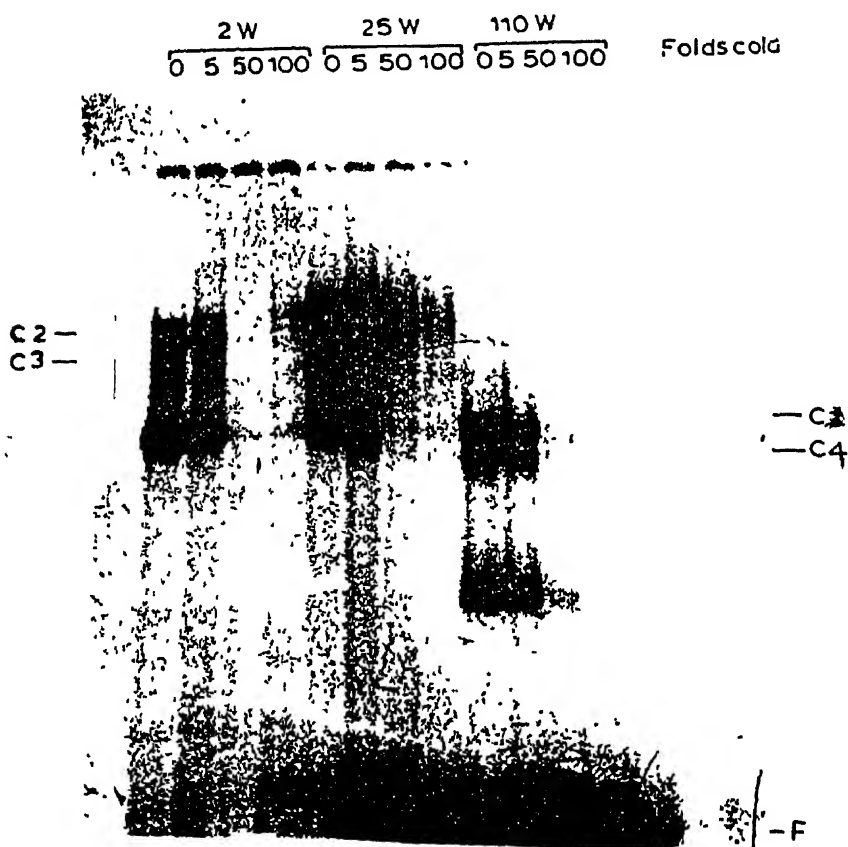


FIG 5 Mobility shift assay for binding of FNT gene to proteins in the nuclear extracts of the liver of 2-, 25-, 110 week old rats.  $^{32}\text{P}$ -labelled 25-mer dsDNA containing the CRE was incubated with 10  $\mu\text{g}$  of nuclear extract, 2  $\mu\text{g}$  of poly dI-dC and different concentrations of cold CRE as mentioned for each lane in reaction mixture. It was resolved on 5% polyacrylamide gel in TBE buffer. C1-C5, different nucleo-protein complexes. F-free DNA (Singh & Kanungo, 1993).

when a DNA fragment gets bound to a protein then its mobility during gel electrophoresis is retarded and is less than that of the unbound DNA. From both the studies it was found that the nuclear extract has more than one *trans*-acting protein factors that bind to the CRE, and that their levels decrease with increasing age. Furthermore, not only their affinities for the CRE change but also new CREBs of different sizes appear to be expressed in old age (Singh & Kanungo, 1993). Thus, the decrease in the expression of the FNT gene during aging is both due to conformational changes that occur in its promoter, and decrease in the levels and affinities of *trans*-acting nuclear factors that are required for regulation of its transcription.

There are certain genes, which on the other hand, show increase in expression as a function of age. We have shown by slot blot hybridization that the level of mRNA for  $\alpha$ -skeletal actin ( $\alpha$ -SKA) of the heart is significantly higher in old age. However, its level does not change in the skeletal muscle.  $\alpha$ -SKA is important for contraction of muscle. The continuous pressure overload of the heart may lead to its higher expression which may contribute to its hypertrophy in old age (Jaiswal & Kanungo, 1990). We have also shown that the expression of certain oncogenes, particularly of *c-myc* and *c-fos*, increases in the heart and brain with increasing age (YK Jaiswal & P.C. Rath unpublished data).

The above studies show that the expression of all genes does not decrease with increasing age. Certain genes decrease in expression, whereas the expression of a few others increase. Though all genes require the same transcription factors for transcription, the regulation of their transcription is brought about by different *trans*-acting factors. In a model proposed to explain the molecular mechanism of aging (Fig. 6; Kanungo, 1975, '80, '93) it has been suggested that reproduction and other stresses such as temperature, pollution, radiation, etc. which an organism encounters during its adult life may deplete certain gene products (*trans*-acting factors) that are necessary for regulation of genes that are vital for normal life. The factors are not replenished. Certain factors may also accumulate and stimulate the expression of some genes including undesirable genes such as oncogenes. These factors are not removed. The *trans*-acting factors modulate/regulate the expression of a gene by binding to *cis*-acting elements located in its promoter. Though *cis*-acting elements and *trans*-acting factors are limited in number, variations in the regulation of genes are brought about through different combinations of their interactions. Thus, their effects are modular in nature. A decrease or increase in one or more *trans*-acting factor may alter the expression of a

gene. Also regulation of a gene by three *trans*-acting factors may be different from that of two factors. The levels of *trans*-acting factors may vary among different individuals of a species when they are subjected to different types of stress. This may affect the expression of genes differently in different individuals of a species. This may account for the variability in the life spans of individuals within a species. Thus, it should be possible to modulate and maintain the expression of certain vital genes by manipulating the levels of their *trans*-acting factors by hormones and other effectors, and thereby extend the period of adulthood (Kanungo, 1993).

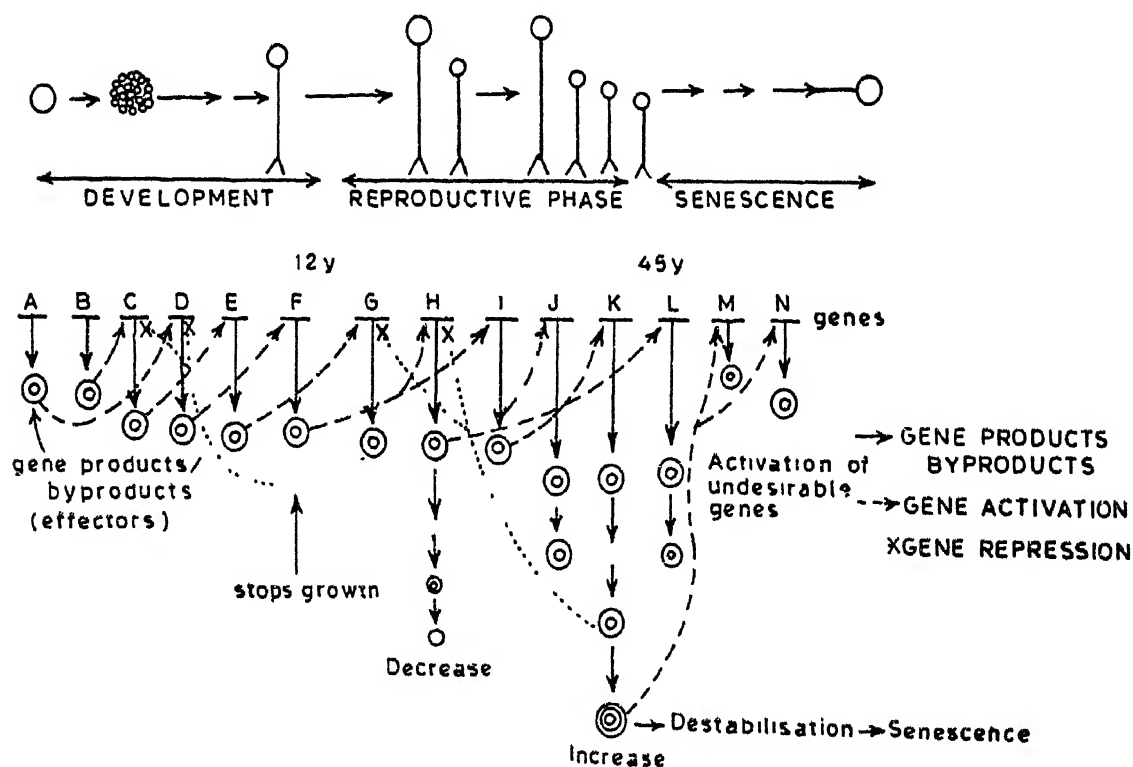


FIG 6 Model for aging (Kanungo, 1980, '93) Upper part - representation of various phases of the life span. Lower part - the number of active genes has been kept at a minimum for clarity. Developmental and reproductive phases are dependent on unique genes, A-F and G-L, respectively. No specific genes for aging are envisaged in this model. Development proceeds by sequential activation of genes A-F through their by-products. Some of the unique gene, E and F, switch on some unique genes, G and H, belonging to early reproductive phase. These then switch on sequentially other genes of the reproductive phase. Continued reproduction causes depletion of certain gene products (*trans*-acting factors) which are essential for keeping some vital genes active. Also some gene products may accumulate and activate certain undesirable genes M and N. Such changes cause destabilisation of homeostatic functioning of genes of reproductive phase leading to aging.

## ACKNOWLEDGEMENTS

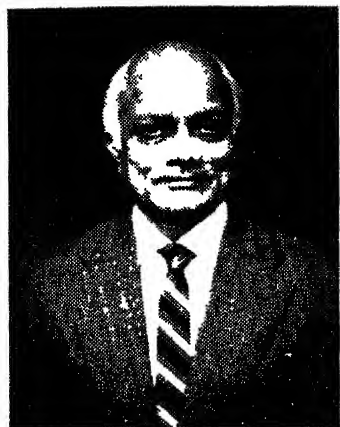
The 1.2 kb cDNA FNT probe was a gift of Prof. R O Hynes, M.I.T., U.S.A. The 1.0 kb cDNA albumin probe was a gift of Dr. T D Sargent, CalTech, U.S.A. The work on the fibronectin gene was carried out by Sanjaya Singh and was supported by Department of Science & Technology, Govt. of India. The work on the albumin gene was carried out by Anita Singh and was supported by Council of Scientific & Industrial Research. I am thankful to Jawaharlal Nehru Memorial Fund for a Jawaharlal Nehru Fellowship, and to the C.S.I.R for an Emeritus Scientistship during which this work was carried out.

## References

- 1 Chaturvedi M M and Kanungo M S Analysis of conformation and function of the chromatin of the brain of young and old rats, *Molec Biol Rep* 10 (1985a) 215-219
2. Chaturvedi M M and Kanungo M S Analysis of chromatin of the brain of young and old rats by nick-translation; *Biochem Biophys, Res, Commun* 127 (1985b) 604-609
- 3 Hynes R O Fibronectins ( New York, Springer-Verlag) (1990)
- 4 Jaiswal Y K and Kanungo M S Expression of actin and myosin heavy chain genes in skeletal, cardiac and uterine muscles of young and old rats, *Biochem Biophys Res Commun*, 168 (1990) 71-77
- 5 James T C and Kanungo M S Effect of sex steroids on cholineacetyltransferase and acetylcholinesterase of cerebral hemisphere of male rats of various ages, *Biochim, Biophys Acta* 538. (1978) 205-211
6. Kanungo M S A model for aging, *J Theo Biol* 53 (1975) 253-261
7. Kanungo M S *Biochemistry of Ageing* (1980) (London: Academic Press)
- 8 Kanungo M S *Genes and Aging* (1993) (Cambridge U.K., Cambridge University Press)
- 9 Kanungo M S and Gandhi B S Induction of malate dehydrogenase isoenzymes in the liver of young and old rats; *Proc Natl Acad Sci U.S.A* 69: (1972) 2035-2038
10. Singh A, Singh S and Kanungo M S Conformation and expression of the albumin gene of young and old rats; *Molec Biol Rep* 14: (1990) 251-254
11. Singh S and Kanungo M S DNase I hypersensitive sites of the 5' region of the fibronectin gene of the liver of the rat; *Biochem Biophys Res. Commun* 181: (1991) 131-137

12. Singh S and Kanungo M S Changes in expression and CRE binding proteins of the fibronectin gene during aging of the rat: *Biochem Biophys Res Commun* **193**: (1993) 440-445
13. Singh S N and Kanungo M S Alterations in lactate dehydrogenase of the brain, heart, skeletal muscle and liver of rats of various ages. *J Biol. Chem* **243**: (1968) 4526-4529





**Holenarasipur Yoganarasimham Mohan Ram** (b. 24 September 1930) did Ph.D. (1959) from University of Delhi. He is Professor of Botany (1968- ), University of Delhi, Delhi; was formerly Fulbright-Smith-Mundt Fellow, Cornell University, USA (1958-60); Unesco Senior Fellow, CNRS, Gif-sur-Yvette, France (1970-71).

The main contributions of Mohan Ram's research group are in : (i) flower development, hormonal control of sex expression and physiology of shelf-life of flowers; (ii) structure, development and reproduction in aquatic angiosperms; (iii) micropropagation of legumes, ornamentals and bamboos; (iv) reproductive biology of trees; (v) structure of wood and gum-resin secreting tissues; and (vi) improvement of techniques of tapping gum resin. He has edited/authored four books; Editor of Publications INSA, (1979-82).

Mohan Ram is Fellow of Indian Academy of Sciences (Vice-President, 1986-88), National Academy of Sciences (India), National Academy of Agricultural Sciences (India); President, International Society of Plant Morphologists (1985-88); he has been Secretary, INSA (1985-88). He is the recipient of J.C. Bose Award (Hari Om Ashram Trust, UGC) (1979), Professor Panchanan Maheshwari Medal (Indian Botanical Society) (1980), President, Botany Section, Indian Science Congress (1980); UGC National Lecturer (1980); Om Prakash Bhasin Award (1986); Sergey G. Navashin Medal (St Petersburg) (1990); Professor S.B. Saxena Memorial Award (INSA) (1990), K.N Bahl Memorial Gold Medal (Society of Biosciences, India) (1992).

*Holenarasipur Yoganarasimham Mohan Ram was elected to the fellowship of the Academy in 1978.*

**Address:** Professor of Botany, University of Delhi, Delhi 110 007; Tel: (011) 7257725/253; Res: 38/4 Probyn Road, Delhi 110 007; Tel: (011) 7257978.

# BIOLOGY OF AQUATIC FLOWERING PLANTS

H Y MOHAN RAM FNA

## INTRODUCTION

Flowering plants have fascinated biologists as they represent the most successful and dominant group of plants on earth. Among these roughly 1% of the species occupy fresh water habitats. In spite of their varied systematic relationships, the aquatic flowering plants share many similar characteristics necessitated by a common habitat. In readapting themselves to life in water they have undergone many structural and functional modifications. In many hydrophytes the plant body has undergone so much reduction that roots may be altogether absent and unusual structures that cannot be clearly assigned to either shoot or leaf categories may be present (Rutishauser & Sattler 1985, 1987 & 1989). The extent of mechanical and vascular tissue development in these plants varies depending on whether they are attached, free-floating, emergent, submerged or amphibious. The system of air-cavities, leaf form, development of special structures for buoyancy, mode of pollination, dispersal and germination of seeds are also diverse.

Arber's (1920) 'Water Plants' is an outstandingly important work as she has discussed their adaptation and evolution. Although there are taxonomic accounts of aquatic plants (Cook *et al.* 1974), comparatively few have been studied in detail, probably because of inconvenience or inaccessibility or even plain apathy. Sculthorpe (1967) has compiled an excellent account of the biology of aquatic vascular plants.

With the unprecedented rise in human population and depletion of resources, there has been a compulsive demand for the study and utilization of freshwater ecosystems. Aquatic plants can be used for food, feed, fibre, paper pulp, green manure, and even for mineral recycling (Boyd 1968, 1970, 1972).

On account of their propensity for adventive spread, aquatic weeds interfere in many of man's activities. In fact, the menace caused by the fast-spreading and pernicious aquatic weeds such as *Eichhornia*, *Salvinia*, *Pistia* and others in blocking waterways and in hampering navigation, irrigation and power generation, is becoming expensive and

unmanageable. Aquatic weeds also harbour mosquitoes and other vectors that carry diseases. Recognizing the high rates of growth of the aquatic weeds and the enormous difficulty in eradicating or even controlling their spread within reasonable limits, there have been attempts to develop methods of utilizing them productively (Anonymous 1976a).

The book "Water-Hyacinth" by Gopal and Sharma (1981) has brought together the enormous amount of literature published until a decade ago, pertaining to morphology, systematics, distribution, general biology, ecology, control and utilization of this most troublesome weed. Jamil (1990), in her recent book, has given a concise account of researches carried out over a decade in the Indian Institute of Chemical Biology (formerly Regional Research Laboratory), Hyderabad, on the environmental biology of water hyacinth. The compilation covers areas such as biology, control strategies, pollution, chemistry, biomass production and bioaccumulation of heavy metals.

Indian flora abounds in a wide variety of aquatic angiosperms (Biswas & Calder 1937, Subramanyam 1962, Deb 1976). The vast potential of aquatic plants has been only marginally utilized by the traditional practices in India and no scientific approach has been developed for deriving greater economic and ecological benefits.

The scope of biological studies on aquatic plants is very vast. Owing to limitation of time, I shall confine my lecture to findings made by our research group using the technique of axenic culture of whole plants. Much of our work is not yet published. However, references would also be made to works of others and areas of interest wherever necessary.

### ESTABLISHMENT OF CULTURES AND NUTRITION

Raising entire plants under precisely controlled and aseptic conditions provides unlimited opportunities to understand the problems of organization, integration and interaction. Aquatic plants are specially advantageous for this purpose because of the ease with which they can be grown and multiplied in liquid culture. The role of tissue culture in the study of the biology of aquatic plants has been emphasized earlier by Mohan Ram (1978) and Mohan Ram and Kakkar (1983). The convenience with which nutrients can be supplied makes the axenic culture of whole hydrophytes ideally suited for understanding chemical ecology, to do critical analytical work dealing with metabolism and to visualize fine structure. Another feature that is commonly observed in submerged

aquatic plants is the absence of a callus, which often becomes problematic in the case of cultured tissues in land plants.

One major obstacle that has dissuaded many investigators from using hydrophytes as experimental material is the inherent difficulty in establishing aseptic cultures. Examination of any part of a water plant shows myriads of associated microorganisms. Secondly, many water plants are highly sensitive to nascent chlorine and other surface-sterilizing agents and succumb to treatments aimed at obtaining aseptic cultures.

The technique recommended by Hillman (1961) for surface-sterilizing the fronds might be ideally suited for duckweeds, especially because they seldom bear fruits in nature. However, our experience with other aquatic plants has been quite different. Undehisced mature fruits and seeds (fresh or dry) constitute a good starting material. A hard pericarp or seed coat protects the embryo from the toxic effects of the sterilizing agents. Successful cultures have thus been raised in *Utricularia* (Mohan Ram & Doreswamy 1966, Mohan Ram & Dutta 1966), *Vallisneria* (Uma & Mohan Ram 1972), *Ceratophyllum* (Mohan Ram & Kapoor 1974, Sehgal 1976) and *Limnophila* (Rao & Mohan Ram 1981), *Neptunia oleracea* (Kakkar & Mohan Ram 1986), *Myriophyllum*, *Trapa* (Mohan Ram & Agrawal, unpublished). Very recently members of the Podostemaceae have also been raised *in vitro* (Vidyashankari & Mohan Ram 1987, Mohan Ram & Sehgal, unpublished work).

Once initial axenic cultures are established, they can be cloned on a defined mineral medium containing sucrose. Experiments designed to find out whether or not an energy source is required for the growth of these chlorophyllous plants in light have demonstrated that without sucrose growth is rather poor (Hillman 1959, Doreswamy & Mohan Ram 1969, Sehgal 1976). In the presence of 2% sucrose, however, the intensity of light required is rather low, suggesting that light probably exercises a morphogenetic role (Hillman 1957). Vitamins do not appear to be necessary, although growth is markedly better in their presence (Doreswamy & Mohan Ram 1969, Uma & Mohan Ram 1972, Sehgal 1976).

The pH of natural waters in which aquatic plants of Delhi grow varies between 7.0 and 10.2 and uptake of iron is problematic. Iron supplied as ferric chloride is not effective in supporting normal growth *in vitro*. The requirement of chelating agents such as EDTA (ethylene-diamine-tetraacetic acid) and EDDHA (ethylene-diamine-di-o-

hydroxyphenylacetic acid) for a large number of water plants cultivated *in vitro* emphasises the possible role of natural chelates in fresh waters (Fogg 1959, Hillman 1959, Doreswamy & Mohan Ram 1969, Gupta & Maheshwari 1970, Venkataraman et al. 1970, Uma & Mohan Ram 1972, Sehgal 1976). The nutritional requirements of water plants are simple. Even growth hormones are not needed or are required in very low concentrations.

### *Bladder Worts*

The insectivores are one of the most curious group of plants endowed with special devices to attract, capture and digest their prey. There have been claims that the nitrogen derived from the insect prey is an obligate requirement for some of these plants (Lloyd 1942, Pringsheim & Pringsheim 1962, 1967, Harder 1963). *Utricularia* is a large carnivorous genus with a predominant number of species occurring in water (Taylor 1989). *Utricularias* are autotrophic but have a supplementary diet ranging from algae, bacteria, and fungi to worms, water insects and other small aquatic organisms (Mohan Ram 1989). Pringsheim and Pringsheim (1967) reported that peptone or beef extract is essential for the cultured plants of *U. minor* and *U. ochroleuca*. These substances promoted good vegetative growth in *U. exoleta* and were absolutely essential for flowering (Harder 1963). In axenic cultures of *U. gibba* (Mohan Ram & Dutta 1966) beef extract and tryptone supported maximum vegetative growth and delayed the abscission of bladders. However, in *U. inflexa* var. *stellaris*, growth occurred even on a medium containing nitrate, thus demonstrating that there is no absolute requirement for organic nitrogen or protein (Doreswamy & Mohan Ram 1969). The presence of organic nitrogen sources such as casein hydrolysate and yeast extract stimulated higher rates of growth. The natural question that arises is: what is the function of the bladders in *Utricularia*? It appears that their development must have been a morphological mutation, which during evolution, must have become secondarily useful in an environment deficient in nitrogen. This view is strengthened by the fact that under the same ecological conditions, numerous other flowering plants exist without any special contrivances for deriving extra sources of nitrogen.

### *The Podostemaceae*

There are several aquatic plants which cannot be grown under aseptic conditions using conventional methods. The present account will be restricted to the family Podostemaceae which has 49 genera and 240

species. Of these 18 species are endemic to India (Nagendran 1975). All Podostemads have an astonishingly reduced thalloid plant body (lacking stems and roots) that resembles a *Fucus*, lichen or a member of the Hepaticopsida. They are found growing in violent, gushing mountain streams, rivers or cataracts and live in submerged conditions attached to rocks and boulders by means of a gum secreted by their rhizoids. Flowering is initiated when the water level recedes. All previous attempts to germinate the seeds of podostemads and raise plants have been unsuccessful. Pannier (1960) has observed that the reason for lack of research on this interesting group of plants is mainly because they cannot be maintained alive out of their native habitat for any period of time. podstemaceae are embryologically unique because they have (i) only syngamy (triple fusion does not occur and therefore endosperm is lacking). (ii) different types of embryo sac, and (iii) a pseudo embryo sac. They are physiologically enigmatic. There is therefore, much to learn by raising plants in pure culture.

When seeds of some members of the Podostemaceae such as *Griffithella hookeriana*, *Indotristicha ramosissima* and *Polypleurum stylosum* were transferred to a nutrient medium aseptically, they failed to germinate. However, using an ingenious technique of placing the seeds on thermocole (polystyrene foam) cubes and floating these cubes on the sterile nutrient medium led to 90% germination. The seed coat secretes a gum which helps in strong adherence to a solid substratum prior to germination. In nature seedlings become established on rocks or wood pieces in rivers. For the sustained growth and development of the podostemads all the constituents of MS (Murashige & Skoog 1962) medium have to be diluted to 1/5 the original concentration (Vidyashankari & Mohan Ram 1987).

## REGENERATION

Probably as an adaptation to turbulence and other mechanical disturbances in water which break up the plant body, the propensity for regeneration and vegetative propagation in the hydrophytes is extraordinarily high. It is not uncommon to see portions of plants and small islands of weeds like the water hyacinth floating in the rivers. Gopal and Sharma (1981) have reviewed the literature on the rates of vegetative propagation in water hyacinth. Two plants are able to multiply to 1200 in 120 days. Relative growth rate and doubling time are probably determined by temperature and nutrient status of water bodies. Fragmentation of the plant body

followed by regeneration from any small part bearing a bud is a common phenomenon. This has been substantiated by *in vitro* studies carried out using *Utricularia*, *Ceratophyllum*, *Limnophila aquatica*, *Neptunia oleracea* and *Myriophyllum oliganthum* in which explants of different sizes have been used (Doreswamy & Mohan Ram 1969, Sehgal 1976, Rao & Mohan Ram 1981, Kakkar & Mohan Ram 1986, Mohan Ram & Agrawal, unpublished work). In *Utricularia inflexa* var. *stellaris* and *Ceratophyllum demersum* excised internodes that lacked buds were incapable of regeneration. In the former, isolated leaf lobes, however, readily gave rise to new plants by initiating adventitious buds (Doreswamy & Mohan Ram 1969).

### HETEROPHYLLY

Formation of several kinds of leaves on the same individual is a characteristic feature of many rooted amphibious angiosperms. This is termed heterophylly. The aerial leaves of an amphibious plant are often morphologically similar to leaves of terrestrial plants and markedly differ from the submerged leaves borne on the same stem in thickness, size, lack of dissection of lamina and in possessing stomata, cuticle etc. Amphibious plants, therefore, exhibit a developmental plasticity in response to a switch between aerial and aquatic medium.

Several environmental factors ( $\text{CO}_2$ , photoperiodism, red, far-red light, growth regulators etc.) have been implicated in determining whether the morphology will be characteristic of the submerged or the aerial leaf. Different concentrations of  $\text{CO}_2$  in air and in water evoke specific responses. For example, when plants of *Ranunculus flabellaris* and *Myriophyllum brasiliense* are grown in a stream of 5%  $\text{CO}_2$  in air on a solid substrate, submerged leaves are formed. When 0.03%  $\text{CO}_2$  is bubbled through water, the plants grow poorly and the leaves are intermediate in form between those produced on land and in water (Bristow 1968). In *Proserpinaca palustris* (Kane & Albert 1982), *Ranunculus flabellaris* (Johnson 1967) and *R. aquatilis* (Cook 1969) short photoperiods induce the submerged, and long photoperiods the aerial leaf form under submersed and emersed conditions. In *Hippuris vulgaris* (Bodkin et al. 1980) and *Marsilea vestita* (Gaudet 1965) high red/far-red ratio promoted the development of the submerged leaf and the converse caused the production of aerial leaf form, irrespective of submergence or emergence. It is hypothesized that light and temperature responses may be mediated by growth regulators.

Heterophylly has been investigated to a considerable extent in our laboratory using the classical material *Limnophila*. In *L. indica* and *L. aquatica* the water leaves are arranged in whorls of 6-12 and are pinnately dissected (Rao 1981). During transition from submerged to aerial condition, leaves of the first node at the water-air interphase dry up. The next three or four nodes show a gradual modification of leaf morphology (Mohan Ram & Rao 1982). However, the leaves that appear later become truly aerial. The number of leaves is reduced from 6 to 2 at each node. Entire, opposite-decussately arranged leaves replace the lobed ones.

Interestingly in *L. indica* flowering is induced only when the shoots bear aerial leaves, probably as response to water stress. When nodal explants of submerged shoots of *L. indica* were cultured in Nitsch's liquid medium containing abscisic acid (ABA,  $10^{-9}$ M- $10^{-6}$ M), they bore regenerated shoots with typical aerial leaves in response to  $10^{-7}$ M and  $10^{-6}$ M ABA, even under submerged conditions. The water leaves were completely suppressed. Flowers were induced precociously by ABA even on submerged nodes (Mohan Ram & Rao 1982). Abscisic acid (ABA) induces floating/aerial leaf form in *Potamogeton nodosus* (Anderson 1978), *Myriophyllum* and *Proserpinaca* (Kane & Albert 1989, Mohan Ram & Agrawal, unpublished work) under submerged conditions.

*Ceratophyllum*, a rootless, submerged macrophyte bears whorls of 5-10 leaves which are twice dichotomised (Sehgal 1976). Occasionally one of the last dichotomies also branches further. With the incorporation of 6-benzylaminopurine (BAP) at  $10^{-4}$ M into the medium it was possible to inhibit dichotomy as well as length of the leaf, resulting in short, once-dichotomised leaves. Cytokinins are, therefore, capable of inducing heteroblastic development even in totally submerged plants.

Heterophylly seems to be a device to maximize photosynthesis in both submerged and aerial habitats. The submerged part makes use of  $\text{HCO}_3^-$  ions in water and aerial leaves have a direct access to gaseous  $\text{CO}_2$  and fix it in a typical PCR (photosynthetic carbon reduction) activity, with a high net photosynthetic rate and a low  $\text{CO}_2$  saturation requirement. This need not be an ubiquitous character of all amphibious plants. Alternatively, it may be argued that aerial shoots could serve primarily as a structural support for flower production, pollination, fruit development and seed dispersal with photosynthesis playing a secondary role. The photosynthetic biochemistry of these plants is little understood.



*Trapa natans* var. *bispinosa*, commonly known as water chestnut is extensively grown in Northern India for its edible seeds. It exhibits not only distinct heterophylly but also prominent heterorhizy (Ghosh 1953, Agrawal & Mohan Ram, unpublished work). The submerged, juvenile leaves are linear and ribbon-like and lack swollen petioles, characteristic of mature, floating, rhomboidal leaves. The roots also show distinct differences in structure and function. The roots which are produced in the beginning are achlorophyllous, unbranched, indeterminate and help in anchorage, whereas those produced later are green, branched, determinate and floating. Some authors are of the opinion that the photosynthetic roots are indeed leaves or a pair of stipules (Hooker 1874, Anonymous 1976a). Experiments are being conducted to understand the basic principles underlying heterophylly and heterorhizy (Agrawal & Mohan Ram, unpublished work).

## FLOWERING

Sculthorpe (1967) has remarked that it is in their reproductive phase that hydrophytes betray their terrestrial ancestry with the greatest clarity. Understandably flowering is under the control of light and temperature. The majority of aquatic angiosperms bear flowers at the surface of water or above. Submerged flowers are noted in members of the families of the Najadales, genera of the Hydrocharitaceae and species of *Callitriche*. Many of the tropical hydrophytes do not flower in cultivation. A few hydrophytes have become adapted to juvenile flowering as strategy to survival (Van Steenis 1957).

In aquatic flowering plants the main thrust in flowering has been in probing the mechanism of floral induction especially using cultures of duckweeds. These miniature monocotyledons present distinct advantages. They multiply rapidly and produce a large, genetically homogeneous, clonal population. They are structurally simple (lack vasculature), with no part of the plant being more than a few cells away from the medium. Being aquatic plants they lend themselves to axenic culture in liquid medium. They are able to grow heterotrophically, often with the requirement of only a few minutes of red light a day to maintain growth (Hillman 1976).

Kandeler (1955) was the first to report that two strains of *Lemna gibba* responded as long-day plants. Hillman (1957) induced flowering in *L. perpusilla* under short days by incorporating EDTA into the medium. It

has now been established that chelates such as EDTA and EDDHA are also essential for flowering in other members of the Lemnaceae (Maheshwari & Chauhan 1963, Maheshwari & Seth 1966, Maheshwari & Gupta 1967, Pieterse *et al* 1970). However, *Wolffia papulifera* does not require EDTA for flowering (Maheshwari & Seth 1965). Among the growth regulators, only cytokinins are able to bring about flowering in the duckweeds under non-inductive conditions—kinetin in *L. perpusilla* (Hillman 1957) and zeatin in *W. microscopica* (Maheshwari & Venkataraman 1966). Surprisingly, ABA induces flowering in *L. perpusilla* (Higham & Smith 1969). Recently flowering has been induced in the duckweeds under non-inductive conditions using salicylic acid, acetyl salicylic acid, tannic acid, benzoic acid, 8'-hydroxyquinoline (a Cu-chelating agent) and *c*-AMP (3', 5' -cyclic monophosphate) (Khurana & Maheshwari, 1978, 1980, 1983a, b, 1984, 1986, Khurana *et al* 1988).

The photosensitivity of the different strains of duckweeds maintained under cultural conditions can be altered by certain substances, so as to reverse the flowering response. For example, in strain G-3 of *L. gibba*, which behaves as a long-day plant, 1-5 mM cupric ion fails to induce flowering under continuous light. Conversely, strain 6746 of *L. perpusilla*, which is a typical short-day plant, continues to flower under these conditions. Maheshwari and Gupta (1967) also observed that in a strain of *L. paucicostata*, addition of excess of iron citrate and EDDHA modified the short-day requirement and caused flowering under long days. Copper is believed to act in some unknown way on the ability of the plant to perceive the photoperiodic stimulus. The chelates seem to remove or immobilise the cupric ion. Promotion of flowering in long days is also exerted by sulphhydryl inhibitors (Takimoto & Tanaka 1973).

The work dealing with the flowering in *Utricularia inflexa* var. *stellaris* has been recently summarized by Mohan Ram (1989). This plant is photosensitive and flowering can be induced by exposing 4-week-old cultures grown in long days (non-inductive) to short-day regimes (Mohan Ram & Doreswamy 1966). Although initiation of inflorescence primordia occurred with 10 cycles of 16 hr dark + 8 hr light or 14 hr dark + 10 hr light, fruit and seed set took place only after 20 or more cycles were given. Isolated shoot tips (measuring 1.0 to 1.5cm in length) were able to perceive the flowering stimulus even when exposed to 7 inductive cycles, as evidenced by the development of float primordia and flower buds (Mohan Ram *et al.* 1972). Organic nitrogen supplied as yeast extract inhibited flowering, while beef extract, tryptone, peptone, casamino acids

and casein hydrolysate depressed the percentage of cultures flowering (Doreswamy & Mohan Ram 1971). When the effects of different sources of iron were studied, maximum flowering was noticed in the medium containing Fe-EDTA (91.6%). Only 8.3 and 4.1% of the cultures flowered on FE-EDDHA and ferric citrate respectively (Doreswamy & Mohan Ram 1971).

Flowering could not be induced by cytokinins under non-inductive conditions. Under short days, kinetin and benzyladenine enhanced flowering and zeatin depressed it. Gibberellic acid ( $GA_3$ ) also lowered the percentage of flowering cultures. A commercial sample of ethephon (2-chloroethyl phosphonic acid) induced flowering in all cultures kept under long days at  $10^{-8}M$ . However, a highly purified sample caused induction but not flower development (Mohan Ram *et al.* 1972).

Pieterse (1985) has recently stated that *Eichhornia crassipes*, one of the 10 world's worst weeds, flowers under both long-day and short-day conditions but detailed studies on the effect of different day lengths on flowering have not been conducted. Plants do not flower below  $21^{\circ}C$ . Gibberellic acid strongly promotes flowering (Pieterse *et al.* 1976).

## POLLINATION

The pollination biology of aquatic angiosperms is a fascinating field of enquiry. In those members in which the flowers are exerted above the level of water, the agencies for pollination could be either air or insects. The elevation of the inflorescence is accomplished by the production of aerial stems, on which the flowers are borne or by the buoyancy provided by rafts of floating leaves as in *Cabomba*, *Nymphoides* and *Hottonia*, and also by whorls of specially designed floats as seen in *Utricularia* (Sculthorpe 1967). It is only the submerged flowers that are truly hydrophilous. Excellent accounts on the mechanism of pollination are available for *Vallisneria spiralis* (Wylie 1917, Kausik 1939) and *Hydrilla verticillata* (Ernst-Schwarzenbach 1945).

Sehgal (1976), working at the University of Delhi has studied pollination in two species of *Ceratophyllum in vitro*. The plants are monoecious. The male flowers bear a large number of spirally arranged stamens. Each stamen has a float at its apex. Owing to a continuous system of lacunae, gases formed elsewhere in the cells are accumulated in the floats. At maturity the basal cells of the stamens give way, and aided by the buoyancy of the floats, the stamens rise to the surface of water. It is

curious that a few pollen grains germinate even within the anthers of intact stamens. Abscission of the stamens stimulates further pollen germination (about 10% in nature and almost 100% *in vitro*). After floating on the surface for a day, the anthers dehisce liberating all the pollen—both ungerminated and germinated. These gradually sink, bringing about pollination on the way. The pollen production is so copious that the bottom of the culture vessel shows a weft of pollen tubes.

It has been confirmed that this method of pollination is not unique only to *Ceratophyllum* plants cultured *in vitro* but it is also true of those growing in nature. The specific advantage of studying pollination *in vitro* lies in the fact that observations can be restricted to a small volume under controlled conditions.

Studies conducted by Mohan Ram and Doreswamy (1966) in *Utricularia inflexa* var. *stellaris* have shown that the flowers undergo self-pollination. In this plant the inflorescence is held aloft by a circlet of floats. In an open flower the stamens are seen closely appressed to the funnel-shaped stigma. The pollen grains germinate *in situ* and the germinated grains are deposited en masse by the inward bending of the stamens and the dehiscence of the anthers. A good seed set has been recorded *in vitro*, emphasizing that the plants are self-compatible. A similar mechanism of pollination was described by Kausik and Raju (1955) in *U. reticulata* under field conditions. Thus self-pollination is of special significance in the reproductive biology of the bladderworts.

Even though pollination in many hydrophytes is accomplished above water, there is often a post-fertilization bending of the peduncle so that the development of the fruits and seeds takes place under water. This has been observed in *Eichhornia crassipes* (Penfound & Earle 1948), *Aponogeton* and other members of the Potamogetonaceae, *Nymphoides*, *Nymphaea* and other Nymphaeaceae members, *Trapa* and *Vallisneria* (Sculthorpe 1967). In *Vallisneria* pollination occurs at the surface of water. After fertilization the peduncle becomes spirally coiled and pulls the developing fruit under water (Kausik 1939). Funke (1938, 1939) reported that auxins like indoleacetic acid and naphthaleneacetic acid stimulate the coiling of the peduncle and also parthenocarpic development of the fruit. Further studies on the hormonal interactions that regulate the pre-pollination elongation and post-fertilization coiling of the peduncle could be meaningfully studied *in vitro*.

## SEED STRUCTURE, VIABILITY AND GERMINATION

Aquatic flowering plants display a wide range of structure in their seeds. Reference will be made only to the work done in our laboratory. Embryos with features of special interest occur in some of them. *Utricularia inflexa* var. *stellaris* is unique in having an unorganized, bun-shaped embryo. On transferring the seeds into liquid White's nutrient medium, 6-14 cotylenoids (a term used by Lloyd 1935, 1942) emerge from the 'plumular pole' in two whorls in 4-6 weeks time. The 'radicular pole' fails to form roots. Since there is no preformed shoot meristem in the embryo, a shoot originates lateral to one of the cotylenoids of the inner whorl. Kumazawa (1967) had made a similar observation in *U. pilosa*. He further demonstrated that by the removal of the shoot, a new shoot was formed at the site of the next cotylenoid of the spiral and he, therefore, believed that a shoot is equivalent to a leaf. All attempts made by us to induce a root in the seedlings of *Utricularia* by growth hormones have failed.

The mature embryo of *Ceratophyllum* is unique among angiosperms in having as many as 12-14 whorls of leaves with a few lateral branches. The oldest lateral branch bears 3-5 whorls of leaves. It may be aptly described as "a complete plant in miniature" (see Guppy 1894). A radicle is absent. It appears that the theory of compensatory growth finds some validity as applied to the situation in *Ceratophyllum*. The embryonic shoot grows at the expense of the root and perhaps also to secure its survival, especially in the absence of roots. Several plant hormones used failed to stimulate meristematic activity in the radicular pole as also noted in the case of *Utricularia*. This finding has confirmed the view that the genetic block to root development cannot be reversed by hormonal treatment.

The seeds of *Trapa natans* are triangular in outline and lack endosperm. Bulk of the seed consists of a single, large, starchy cotyledon. The second cotyledon is rudimentary and is present on the hypocotyl. A radicle is absent in the embryo. Germination of the seed occurs by the growth and elongation of the "mesocotyl" (Philomena & Shah 1985), which carries the plumule, the smaller cotyledon and hypocotyl with it. The hypocotyl also grows in a negatively geotropic manner, and gives rise to several lateral roots, which help in anchoring the seedling. The plumule then begins to grow up, ascending in water till the shoot apex reaches the surface of water where the apical rosette of leaves is formed.

*Griffithella hookeriana*, a member of the Podostemaceae has two large cotyledons, a radicle and a hypocotyl. However, a discernible plumule is absent. On germination, the radicular pole forms unicellular rhizoids (often forked at the tip) instead of roots. The plumular pole remains quiescent for about 8 days after which a tiny leaf primordium appears between the two cotyledons. The formation of shoot is, therefore, post-germinal. The shoot apex forms 2 or 3 pairs of leaf-like outgrowths and then ceases to function. A protuberance which is initially small, thin, green, and round, oval or elongate is formed from the primary axis below the cotyledons. It gradually expands into a horizontal thallus (Vidyashankari & Mohan Ram 1987). The study of germination and establishment of plants in several other members of the Podostemaceae is in progress in the lecturer's laboratory.

Seeds of most hydrophytes have a prolonged dormancy. Some remain viable only for a few months. However, those of *Nuphar* and *Nymphaea* retain their viability even after being frozen in ice or mud for several weeks. A 237 year old seed of Asiatic *Nelumbo nucifera* germinated on a herbarium sheet after accidental flooding in the British Museum of Natural History, when firemen were trying to put out flames due to bombing in World War II (Ramsbottom 1942). Ohga (1926) had collected sample of ancient seeds of *N. nucifera* from a dried lake-bed in Manchuria. As nearly all of them could be germinated by treatment with sulphuric acid, the question of their true age became an enigma. Various estimates ranging from 120 to 1040 years were suggested (Libby 1951). Using the same material Godwin and Willis (1964) estimated an age of not more than 100 years (Sculthorpe 1967).

### SOMACLONAL VARIATION

In the cultures of *Utricularia inflexa* var. *stellaris* raised from shoot tip or stolon explants (from plants kept under long days which ensures vegetative condition) growth of a markedly different kind was observed. In contrast to the normal habit, fine thread-like stolons with profuse branches with shorter and sparingly dissected leaves were noted. This abnormal condition was termed "bushy clone" (Mohan Ram *et al*, 1972). The bushy habit was maintained in subculture. The other characteristics of these bushy plants were the failure to flower under any photoperiodic treatment (the normal plants flower under short days) and the tendency to sink. In nature, plants showing this habit have not been recorded. It is a variation that has developed as a result of tissue culture.

An abnormal type of "bushy seedling" was also observed in the cultures of *Ceratophyllum demersum* and *C. echinatum* along with normal plants (Sehgal 1976). Bushy seedlings were produced both from excised embryos and embryos enclosed in fruits, collected from plants growing *in vivo* and *in vitro*. There was a higher incidence of these abnormal seedlings in immature fruits, or those produced early in the season. They were characterized by the absence of chlorophyll, slow growth rate, distorted fleshy cotyledons, condensed internodes, short leaves and small white shoot apices. Excised portions of these seedlings failed to regenerate. High (36°C) and low (6°C) temperature treatments, and growth regulators such as kinetin, benzyladenine, GA<sub>3</sub>, and NAA could not induce normal growth. Incorporation of growth retardants, TIBA (triiodobenzoic acid), 2, 4, 5-T (2, 4, 5-trichloroacetic acid) and chlorflurenol failed to induce the morphological features unique to the bushy habit in normal plants. Surprisingly, some of the bushy seedlings reverted to normal growth by the production of a normal axillary bud or by transformation of the shoot apex. Once transformed, the normal branches even produced flowers and fruits.

Such bushy plants have never been observed in nature, even though the massive cotyledons could have supported their growth for a short period. It is likely that severe competition would have eliminated these plants from nature and *in vitro* culture sustains all such plants which have no chance of survival in nature.

### ECONOMIC AND UTILIZATION ASPECTS

A brief reference was made earlier in the lecture to the immense problems created by the water weeds and the amount of money and effort invested to eradicate them with little success. The chemical, mechanical and biological methods would be continued to be maximally used to combat the menace caused by aquatic weeds. Majid (1986) has given a detailed account of the utility and development of aquatic weeds. Abbasi et al. (1988) have reviewed the distribution, impact and control of aquatic weeds. No attempt will be made here to go into those aspects in detail. It is increasingly felt that removing aquatic weeds and selling or utilizing them to defray the cost of their removal has been an appealing concept. There are numerous other non-weedy hydrophytes which have several positive assets that can be of much utility.

### *Animal Feed*

Aquatic weeds constitute a ready-made crop of much value to wide variety of aquatic fauna and domestic animals as they require no tillage, fertilizer, protection or cultivation practices (Anonymous 1976a). Aquatic plants can be exploited for animal feed provided they are harvested only from unpolluted waters (Abbasi *et al.* 1988). Grass carp whose meat is highly priced, devours a wide variety of succulent submerged weeds. Grayfish—a relative of the lobster—survives on aquatic weeds in rice fields. Duck, geese and swans clear large areas of aquatic weeds and in turn provide meat and eggs. *Bubalis bubalis* (water buffalo) feeds on submerged vegetation. In Indonesia, water hyacinth is harvested for pig food.

### *Human food*

Many food plants are grown in the world's fresh waters, but none other than rice has attained importance as a staple crop. Although several aquatic plants are cultivated, a few commonly grown in India are mentioned below:

- (a) *Colocasia esculenta* (Taro)—Tuberous roots are used as rice or potato substitute.
- (b) *Nelumbo nucifera* (Lotus)—It is considered sacred for its flowers. Its seeds and rhizomes are eaten in numerous ways in India, China, Thailand, etc.
- (c) *Trapa natans* var *bispinosa* (water chestnut or singhara)—It yields large, starchy, white seeds which are eaten raw, boiled or roasted. The ground starch is also used as a substitute for cereals.
- (d) *Euryale ferox* (Fox nut, gorgon nut or makhana)—This resembles lotus but is of a larger size and bears prickly fruits, whose seeds are roasted and eaten. On roasting, the seed coat swells and bursts and can be easily peeled off.
- (e) *Neptunia oleracea*—Its crisp young twigs are eaten raw or cooked as greens.
- (f) *Ipomoea aquatica* (Water spinach)—Its leaves and stems are used as vegetable.



### Soil Additives

The importance of weeds as agricultural composts and as a source of protein has been stressed in recent years. *Aeschynomene indica* and *A. aspera*, *Sesbania rostrata* and *Neptunia oleracea* which harbour *Rhizobium* fix an enormous amount of nitrogen and when grown with rice or used as soil additives can contribute much for enriching wet soils. Water hyacinth has also been used as compost and sold to farmers to assess its manurial value (Mukherjee et al. 1964). A higher yield of tomatoes and okra has been obtained by using compost of *Pistia*, *Najas*, *Hydrilla* and *Ottelia*.

Although the emphasis in this lecture has been mainly on angiosperms, it may be recalled that the use of *Azolla* as a green manure has been a beneficial practice in southeast Asia. In Vietnam rice yields are 50% higher than normal by cultivating *A. pinnata* in the fields. A blue-green alga, *Anabaena azollae* occurs in the pockets of *Azolla* leaves and fixes atmospheric nitrogen.

### Biofilters

Due to prolific growth rate aquatic weeds can survive and grow on waters containing high BOD and toxic chemicals (Dinges 1976). They take up toxic chemicals and accumulate them in their own system. Water hyacinth has been explored for treating waste waters from dairies, piggeries, textile mills, oil mills and metal work industries (Abbasi 1987).

Water hyacinth, a successful survivor of polluted waters accumulates a large number of heavy metals from its environment. Several reports have appeared on the removal of metals such as copper (Sutton et al. 1971)\*, silver, cobalt, strontium (Wolverton et al. 1975a)\*, mercury, lead (Wolverton et al. 1975b)\*, cadmium and nickel (Wolverton et al. 1975c)\*, by water hyacinth from contaminated effluents. Muramoto and Oki (1983) reported the ability of water hyacinth to remove cadmium, lead and mercury from the metal-containing solution without nutrients in water. This, combined with the ability of the plant to adapt to variety of climatic and environmental conditions makes it ideally suited for metal removal systems.

Jamil (1990 and references cited therein) has presented data on the concentration of heavy metals such as Cd, Hg and Zn by water hyacinth

---

\* See Muramoto and Oki 1983.

from solutions containing cadmium chloride, mercuric chloride and zinc sulphate at 25, 50, 75 and 100 ppm, after a period of 1, 3, 4 and 7 days. In most cases maximum concentration was registered after treatment for 24 hr. In another study water hyacinth plants were treated with toxic metals such as Hg, Cd, Pb and Cr at concentrations ranging from 25 to 100 ppm. The maximum uptake of these toxic metals was found at 25 ppm and the uptake of these toxic metals was found at 25 ppm and the uptake of metals decreased with increase in the concentration of the metal ions. Pb and Cd were absorbed in much larger quantities than Hg and Mn.

Various ecological groups of plants have been studied for their ability to cycle heavy metals from a pond with post-sewage water containing Fe, Mn, Zn, Cr, Cu, Ni and Pb (Ozimek 1985). Of the plant groups, pleustonic plants (comprising *Lemna minor* and *L. gibba*) showed a higher efficiency for removal of metals as they grew well in polluted waters, drew the metals from water, were not limited by light and did not set the sediment in motion. Plants such as *Lemna*, *Myriophyllum*, *Scirpus* and *Pistia* have been used in the treatment of industrial waste waters and nutrient-rich agricultural drainage effluents.

The mechanism used by heavy metal-tolerant plants to alleviate the stresses imposed by the heavy metals is synthesis of heavy metal binding polypeptides. These compounds have been referred to as cadystins and phytochelatins. The enzyme that catalyzes their biosynthesis, phytochelatin synthase, is constitutively expressed in plants. These polypeptides apparently sequester and detoxify excess metal ions and play a role in homeostasis. The nature and role of these polypeptides, which possess a general structure  $(\gamma\text{-Glu-Cys})_n\text{-Gly}$ , where  $n=2-7$ , have been recently reviewed by Steffens (1990).

The study of heavy metal uptake can be critically improved by using axenic cultures and a beginning in this direction has been made in the lecturer's laboratory.

### Energy

Several well-established technologies are available for converting biomass into energy. Thermal, thermochemical and aerobic fermentation are not suitable for aquatic biomass owing to very high water content (90-95%). Drying up of plants such as *Eichhornia* and *Salvinia* to a level where they can be used in conversion processes is expensive. Technological difficulties make anaerobic digestion also unfeasible.

However, if digesters are designed properly, then freshwater as well as marine biomass can be aerobically converted into energy (Abbasi et al. 1988).

#### *Source of paper and pulp*

The fibres of some of the aquatic weeds, notably water hyacinth and *Salvinia* have been used for making paper. The paper passes all the tests with satisfactory results (Abbasi et al. 1988). The other weeds that have been tested as raw materials are *Pistia* and *Typha*.

#### *Ornamental, Decorative and Aquarium Plants*

Aquatic plants such as *Nymphaea* and lotus have been held in very high esteem in India, China and Japan and they appear as motifs in arts, crafts and architecture. Although water gardens are not very popular, several species of *Cyperus*, *Typha*, *Aponogeton*, *Sagittaria*, *Nelumbo*, *Nymphaea* etc. are grown in Indian gardens. Botanical gardens draw huge crowds to witness the massive-leaved *Victoria amazonica*. Although the shola plant, *Aeschynomene aspera* is not grown for beauty but its soft pith-like wood (secondary xylem) is used for decorative items such as flowers, garlands, toys, hats, replicas of temples, lining of palanquin tops, life-belts, swimming jackets etc. (Anonymous 1985). A very large number of craftsmen in India make their living on the shola plant.

The sterile culture and sale of aquarium plants like *Cabomba*, *Vallisneria* and *Brasenia* are a part of horticultural activity in Europe. Aquaria also make use of herbivore fish culture. Fishes of different colours and shapes are especially used for this purpose. Display of aquaria in homes, museums and public institutions is not only of aesthetic value but also draws attention to the need for conservation of fragile, aquatic ecosystems and endangered plants.

### CONCLUSION

In the above account, I have tried to outline the immensely interesting life and activities of aquatic flowering plants and their economic potentiality.

Knowledge of the biology of aquatic weeds has been built largely on observation. However, valuable this approach has been, the causative aspects can be understood only by experimentation and analysis. In addition to summarizing the work already done using *in vitro* culture in

my laboratory, I have suggested a few areas, in which future strides may be intellectually rewarding and practically beneficial.

### ACKNOWLEDGEMENTS

I am indebted to the Indian National Science Academy for asking me to deliver the first lecture in the series instituted to honour Professor Shyam Bahadur Saksena. I have pleasure in thanking Dr (Mrs) Anita Sehgal, Miss Anuradha Agrawal, Ms Anju Khanijo, Mr. A. Saadi and Mrs. G.S. Sundari Iyer for their able and willing help in preparing the manuscript.

### References

1. Abbasi S A 1987 Cited in *Proceedings of Aquatic Macrophytes for Water Treatment and Resource Recovery*, eds W H Smith and K R Reddy (Orlando: Magnolia Publishing) 175pp.  
—, Nipanay P C and Soni R 1988 Aquatic weeds—Distribution, impact and control; *J scient Ind Res* 47 650-661.
2. Anderson L W J 1978 Absciscic acid induces formation of floating leaves in the heterophyllous aquatic angiosperm *Potamogeton nodosus*; *Science N Y* 201 1135-1138
3. Anonymous 1976a *Making Aquatic Weeds Useful. Some Perspectives for Developing Countries* (Washington D C: National Academy of Sciences) 171 pp.  
— 1976b *The Wealth of India: A Dictionary of Indian Raw Materials and Industrial Products—Raw Materials Series*. Vol.X (New Delhi: Publications and Information Directorate, CSIR).  
— 1985 *The Wealth of India: A Dictionary of Indian Raw Materials and Industrial Products* Vol IA (Revised) (New Delhi: Publications and information Directorate, CSIR) 93 pp.
4. Arber A 1920 *Water Plants—A Study of Aquatic Angiosperms* (Cambridge: University Press) 436 pp.
5. Biswas K and Calder C C 1937 *Handbook of Common Water and Marsh Plants of India and Burma, 1936* Health Bulletin No. 24 (Delhi: Manager of Publications) 140pp.
6. Bodkin P D, Spence D and Weeks D 1980 Photoreversible control of heterophylly in *Hippuris vulgaris* L.; *New Phytol.* 84 533-542.
7. Boyd C D 1968 Fresh water plants: A potential source of protein; *Econ. Bot.* 22 359-368.

- 1970 Vascular aquatic plants for mineral nutrient removal from polluted waters; *Econ. Bot.* **24** 1-95.
- 1972 A bibliography of interest in the utilization of vascular aquatic plants; *Econ. Bot.* **26** 74-84.
8. Bristow J M 1968 The effects of carbon dioxide on the growth and development of amphibious plants; *Can. J. Bot.* **47** 1803-1807.
9. Cook C D K 1969 On the Determination of Leaf form in *Ranunculus aquatilis*, *New Phytol* **69** 469-480.  
—, Gut B J, Rix E M, Schneller J and Seitz M 1974 *Water plants of the World –A Manual for the Identification of the Genera of Freshwater Macrophytes* (The Hague: Dr W Junk B.V. Publishers) 477 pp.
10. Deb D B 1976 A study on the aquatic vascular plants of India; *Bull Bot Soc Bengal* **29** 155-170.
11. Dinges R 1976 Water hyacinth cultures for waste water treatment Texas Department of Health, Texas, Austin) 143 pp.
12. Doreswamy R and Mohan Ram H Y 1969 Studies on growth and flowering in axenic cultures of insectivorous plants. I. Seed germination and establishment of cultures in *Utricularia inflexa* Forsk.; *Phytomorphology* **19** 363-371.  
— and — 1971 Studies on the growth and flowering in axenic cultures of insectivorous plants.II. Induction of flowering and development of flowers in *Utricularia inflexa*; *Z Pflphysiol.* **65** 315-325.
13. Ernst-Schwarzenbach M 1945 Zur Blütenbiologie einigen Hydrochantaceen; *Ber Schweiz bot Ges.* **55** 33-69.
14. Fogg G E 1959 Dissolved organic matter in oceans and lakes; *New Biol* **29** 31-48
15. Funke G L 1938 Observations on the growth of water plants. II; *Biol. Jahrb.* **5** 382-403.  
— 1939 Observations on the growth of water plants. III; *Biol. Jahrb.* **6** 334-350.
16. Gaudet J J 1965 The effect of various environmental factors on the leaf form of the aquatic fern *Marsilea vestita*, *Physiologia Pl* **18** 674-686.
17. Ghosh M 1953 *The Vegetative Anatomy, Flower and Fruit of Trapa bispinosa* Roxb.; M.Sc. Thesis. Univ. Delhi, Delhi, India. 45 pp.
18. Godwin H and Willis E H 1964 The viability of lotus seeds (*Nelumbium nucifera* Gaertn.); *New Phytol.* **63** 410-412.
19. Gopal B and Sharma K P 1981 Water-Hyacinth (*Eichhornia crassipes*) the Most Troublsome Weed of the World (Delhi: Hindasia Publishers) 128 pp.
20. Guppy H B 1894 Water plants and their ways. *Ceratophyllum demersum*; *Science Gossip* **1** 195-197.

21. Gupta S and Maheswari S C 1970 Growth and flowering of *Lemna paucicostata*. I. General aspects and role of chelating agents in flowering; *Pl. Cell Physiol* **11** 83-95.
22. Harder R 1963 Blütenbildung durch tierische zusatznahrung und andere Faktoren bei *Utricularia exoleta* R. Braun.; *Planta* **59** 459-471.
23. Higham B M and Smith H 1969 The induction of flowering by abscisic acid in *Lemna perpusilla* 6746; *Life Sci.* **8** 1061-1065.
24. Hillman W S 1957 Photoperiodic control of flowering in *Lemna perpusilla*, *Nature* **181** 1275.  
 — 1959 Experimental control of flowering in *Lemna* I General methods Photoperiodism in *L. perpusilla* 6746; *Am. J. Bot.* **46** 466-473  
 — 1961 The Lemnaceae or duckweeds. A review of the descriptive and experimental literature; *Bot. Rev.* **27** 221-281.  
 — 1976 Calibrating Duckweeds: Light, Clocks, Metabolism, Flowering; *Science* **193** 453-458.
25. Jamil K 1990 Environmental Biology of Water Hyacinth (Vallabh Vidyanagar. Avichal Science Foundation) 104 pp.
26. Johnson M P 1967 Temperature dependent leaf morphogenesis in *Ranunculus flabellaris*, *Nature* **214** 1354
27. Kakkar M and Mohan Ram H Y 1986 Regeneration of whole plants from tissue cultures of the tropical aquatic legume *Neptunia oleracea*, *J. Pl. Physiol.* **126** 83-91.
28. Kandeler R 1955 Über die Blütenbildung bei *Lemna gibba* L. I Kulturbedingungen und Tageslängenabhängigkeit; *Zeit. Bot.* **43** 36-71.
29. Kane M E and Albert L S 1982 Environmental and growth regulator affects on heterophylly and growth of *Proserpinaca intermedia* Mack. (Haloragaceae); *Aquat. Bot.* **13** 78-85.  
 — and — 1989 Absciscic acid induction of aerial leaf development in *Myriophyllum* and *Proserpinaca* species cultured *in vitro*; *J. Aquat Plant Manage.* **27** 102-111.
30. Kausik S B 1939 Pollination and its influence on the behaviour of the pistillate flower in *Vallisneria spiralis*; *Am. J. Bot.* **26** 207-211.  
 — and Raju M V S 1955 A contribution to the floral morphology and embryology of *Utricularia reticulata* Smith; *Proc. Indian Acad. Sci.* **41** 155-166.
31. Khurana J P and Maheshwari S C 1978 Induction of flowering in *Lemna paucicostata* by salicylic acid; *Pl. Sci. Letters* **12** 127-131.  
 — and — 1980 Some effects of salicylic acid on growth and flowering in *Spirodela polyrrhiza* SP<sub>20</sub>; *Pl. Cell Physiol* **21** 923-927.

- and — 1983a Floral induction in *Wolffia microscopica* by salicylic acid and related compounds under non-inductive long days; *Pl. Cell Physiol.* **24** 907-912.
- and — 1983b Effect of 8-Hydroxyquinoline on flowering and endogenous levels of iron and copper in *Lemna paucicostata* strain LP6; *Pl. Cell Physiol.* **24** 1251-1254.
- and — 1984 Floral induction in short-day *Lemna paucicostata* 6746 by 8-hydroxyquinoline, under long day; *Pl. Cell physiol.* **25** 77-83.
- and — 1986 A comparison of the effects of chelates, salicylic acid and benzoic acid on growth and flowering of *Spirodela polyrrhiza*, *Pl. Cell Physiol.* **27** 919-924.
- , Tamot B K and Maheswari S C 1988. Role of Adenosine 3'5'-cyclic monophosphate in flowering of a short day duckweed *Lemna paucicostata* 6746; *Pl Cell Physiol* **29** 1023-1028.
32. Kumazawa M 1967 An experimental study on the seedling of *Utricularia pilosa* Makino, *Phytomorphology* **17** 494-498.
33. Libby W F 1951 Radiocarbon dates.II; *Science* **114** 291-296.
34. Llyod F E 1935 *Utricularia*, *Biol Rev.* **10** 72-110
- 1942 *The Carnivorous Plants* (Waltham: Chronica Botanica Co.)
35. Maheshwari S C and Chauhan O S 1963 In *vitro* control of flowering in *Wolffia microscopica*, *Nature* **198** 99-100.
- and Gupta S 1967 induction of flowering in *Lemna paucicostata*, a short day plant, by chelating agents and iron, *Planta* **77** 95-98.
- and Seth P N 1965 Photoperiodic control of flowering in *Wolffia papulifera*; *Pl Cell Physiol.* **7** 163-165.
- and — 1966 Induction of flowering in *Wolffia microscopica* by iron salt of ethylenediamine-di-o-hydroxyphenylacetic acid (Fe EDDHA). *Z. Pflphysiol.* **55** 89-91.
- and Venkataraman R 1966 Induction of flowering in a duckweed—*Wolffia microscopica*—by a new Kinin, Zeatin; *Planta* **70** 304-306.
36. Majid F Z 1986 *Aquatic Weeds—Utility and Development* (India: Agro Botanical Publishers) pp 49-53
37. Mohan Ram H Y 1978 *In vitro* culture of aquatic Flowering plants: Achievements and prospects, 10th Philip R, White Mem Lec, Ahmedabad 17pp.
- 1989 *Utricularia inflexa* var. *stellaris*, in *CRC Handbook of Flowering* Vol. VI pp 655-660 ed A H Halevy (Boca Raton, Florida; CRC Press).
- and Doreswamy R 1966 Growth and Flowering of *Utricularia inflexa* Forsk. var. *stellaris* Taylor in axenic culture; *Naturewissenschaften* **53** 387.

- and Dutta S 1966 *In vitro* culture of *Utricularia*; *Curr Sci.* 35 48-50.
- , Harada H and Nitsch J P 1972 Studies on the growth and flowering in axenic culture of insectivorous plants. III. Effects of photoperiod, ethrel, morphactin and a few other growth substances and metabolic inhibitors on *Utricularia inflexa*, *Z. Pflphysiol.* 68 235-253
- and Kakkar M 1983 Role of tissue culture in the study of aquatic plants; *Bull Bot Surv India* 25 26-34.
- and Kapoor A 1974 *In vitro* growth and development of horn wort; in 3rd Int. Conf. Plant Tissue and Cell Culture Leicester, UK.
- and Rao S 1982 *In Vitro* induction of aerial leaves and of precocious flowering in submerged shoots of *Limnophila indica* by abscisic acid; *Planta* 155 521-523.
38. Mukherjee R K, Bhanja A, Roy Burman P and Sircard S M 1964 Presence of bound auxin in the roots of water hyacinth (*Eichhornia crassipes*), *Bull. Bot soc Bengal* 18 87-90.
39. Muramoto S and Oki Y 1983 Removal of some heavy metals from polluted waters by water hyacinth (*Eichhornia crassipes*); *Bull Environm Contam. Toxicol.* 30 170-177.
40. Murashige J and Skoog F 1962 A revised medium for rapid growth and bioassays with tobacco tissue cultures; *Physiologia Pl.* 15 473-497.
41. Nagendran C R 1975 Studies on Podostemaceae; Ph.D. Thesis, Univ Mysore, Mysore 173 pp
42. Ohga I 1926 The germination of century-old and recently harvested Indian lotus fruits, with special reference to the effect of oxygen supply; *Am J Bot.* 13 754-759.
43. Ozimek T 1985 Heavy metal content in macrophytes from ponds supplied with post-sewage water; *Symposia Biologica Hungarica* 29 41-50.
44. Penfound W T and Earle T T 1948 The Biology of the water hyacinth; *Ecol. Monogr.* 18 447-472.
45. Philomena P A and Shah C K 1985 Unusual germination and seedling development in two monocotyledonous dicotyledons; *Proc. Indian Acad Sci. (Pl. Sci.)* 95 221-225.
46. Pieterse A H 1985 *Eichhornia crassipes* in *CRC Handbook of Flowering* Vol. 2 462-465 pp ed A H Halevy (Boca Raton, Florida: C R C Press, Inc.).
- , Aris J J A M and Butter M E 1976 Inhibition of float formation in water hyacinth by gibberellic acid; *Nature* 260 388-389.
- , Bhalla P R and Sabharwal P S 1970 Induction of flowering in *Lemna gibba* G-3 by Fe EDDHA, *Pl Cell Physiol.* 11 675-676.



47. Pannier F 1960 Physiological responses of Peltostemaceae in their natural habitat; *Int. Rev. Ges. Hydrobiol.* **45** 347-354.
48. Pringsheim E G and Pringsheim O 1962 Axenic culture of *Utricularia*, *Am. J. Bot.* **49** 898-901.  
— and — 1967 Kleiner Beitrag zur Physiologie von *Utricularia*, *Z. Pflphysiol.* **57** 1-10.
49. Ramsbottom J 1942 Recent work on germination; *Nature* **149** 658.
50. Rao S 1981 *Growth and Morphogenesis in Limnophila*, Ph.D Thesis, Univ. Delhi, Delhi, India, 216 pp.  
— and Mohan Ram H Y 1981 Regeneration of whole plants from the cultured root tips of *Limnophila indica*, *Can. J. Bot.* **59** 969-973
51. Rutishauser R and Sattler R 1985 Complementarity and heuristic value of contrasting models in structural botany. I. General considerations, *Bot. Jahrb. Syst.* **107** 415-455.  
— and — 1987 Complementarity and heuristic value of contrasting models in structural botany II Case study on leaf whorls, *Equisetum* and *Ceratophyllum*, *Bot. Jahrb. Syst.* **109** 227-255.  
— and — 1989 Complementarity and heuristic value of contrasting models in structural botany. III. Case study on shoot like "leaves" and leaf-like "shoots" in *Utricularia macrorhiza* and *U. purpurea* (Lentibulariaceae), *Bot. Jahrb. Syst.* **111** 121-137.
52. Sculthorpe C D 1967 *The Biology of Aquatic Vascular Plants* (London: Edward Arnold) 610 pp.
53. Sehgal A 1976 *Some aspects of Developmental Biology of Ceratophyllum*; Ph.D. Thesis, University of Delhi, Delhi, India. 199 pp.
54. Steffens J C 1990 The heavy metal-binding peptides of plants; *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **41** 553-575.
55. Subramanyam K 1962 *Aquatic Angiosperms Botanical Monograph No.3* (New Delhi: Council of Scientific and Industrial Research) 190 pp.
56. Takimoto A and Tanaka O 1973 Effects of some SH-inhibitors and EDTA on flowering in *Lemna perpusilla* 6746; *Pl. Cell Physiol.* **14** 1133-1139.
57. Taylor P 1989 The genus *Utricularia*—a taxonomic monograph, *Kew. Bull. Addit. Ser.* **14** 1-724 (London: HMSO).
58. Uma M C and Mohan Ram H Y 1972 *In vitro* culture of *Vallisneria spiralis*; *Phytomorphology* **22** 121-124.
59. Van Steenis C G G J 1957 Specific and infraspecific delimitation: *Flora Malesiana. Ser.* **15** 167-236.

60. Venkataraman R, Seth P N and Maheshwari S C 1970 Studies on the growth and flowering of a short-day plant *Wolffia microscopica*. I. General aspects and induction of flowering by cytokinins; *Z Pflphysiol.* **17** 721-731.
61. Vidyashankari B and Mohan Ram H Y 1987 *In vitro* germination and origin of thallus in *Griffithella hookeriana* (Podostemaceae); *Aquat. Bot* **28** 161-169.
62. Wylie R B 1917 The pollination of *Vallisneria spiralis*; *Bot. Gaz.* **63** 135-145.



**Narayana Balakrishnan Nair** (b. 6 July 1927) did Ph.D.(1955), D.Sc.(1965) from University of Madras. He is CSIR Emeritus Scientist (1991- ), and was formerly Dean, Faculty of Science (1976-78) and Head, Department of Aquatic Biology and Fisheries at University of Kerala (1968-80)

Nair has brought to light the taxonomy, functional morphology, embryology, reproductive biology, digestive physiology and ecology of organisms concerned with biodegradation of cellulose in aquatic habitats. Elucidated the biology of several species of shellfish and fin fish, their pathology and aquacultural techniques. Made noteworthy contributions to the ecology of tropical wetlands and their conservation; and ecology of marine algae, mangrove and sea grasses. He is considered an expert on biodeterioration of cellulose materials in aquatic environment. He has Authored three books 'Text Book of Marine Ecology' (MacMillan), 'Discovery of the Ocean' (NCERT), 'Paristhuthi Vignanam' (Ecology) (Govt. of Kerala).

Nair is Fellow of Zoological Society of London, Indian Academy of Sciences, National Academy of Sciences (India), Zoological Society of India, and Marine Biological Association of India; Founder Fellow, National Academy of Agricultural Sciences (India); Member, INSA Council (1978- 80). He is the recipient of Chandrakala Hora Memorial Medal (INSA) (1988), Jawaharlal Nehru Fellowship (1982); Shanti Swarup Bhatnagar Prize (1971); UGC National Lecturer (1979); All India Congress of Zoology Gold Medal (1980), Padma Shri (1984)

*Narayana Balakrishnan Nair was elected to the fellowship of the Academy in 1975.*

**Address:** CSIR Emeritus Scientist, Department of Aquatic Biology and Fisheries, Beach P.O. Trivandrum 695 007; Tel: (0471) 450138; Res: 'Swathi', Residency Road, Thycaud, Trivandrum 695 014, Tel: (0471) 63746.

# **THE PROBLEM OF BIODETERIORATION ALONG THE INDIAN COASTS AND ITS IMPACT ON FISHERIES**

**N BALAKRISHNAN NAIR FNA**

## **INTRODUCTION**

There has probably never been a period in the history of developing countries when interest in the utilisation of timber and progress in this field have been greater than at present. Conservation and protection have, therefore, become extremely essential for the effective utilisation of the limited resources. There are two approaches to the problem of the destruction of wooden structures in seawater exposures. One is discarding the home-grown, susceptible timbers as structural materials for marine constructions and seagoing craft, and using more expensive materials such as steel, concrete; etc.; but even these are not exempt from the ravages of all types of deterioration and are vulnerable to one or more of the destructive processes. The utilisation of steel and concrete for all kinds of marine constructions will certainly be impracticable for a long time in developing countries like India. The other approach is the use of home-grown timber as building material, with the use of every device and technique calculated to prolong their service life. Careful conservation and scientific protection are integral parts of this reasonably sound approach. This involves among others a precise understanding of the biology of the organisms which are responsible for the destruction. Successful control measures depend upon a knowledge of the nature of organisms against which the control is directed. Studies along these lines have been in progress in the University of Kerala.

## **THE ORGANISMS THAT CAUSE DESTRUCTION OF TIMBER ALONG THE COASTS OF INDIA**

Timber in seawater exposures is threatened by decay through the action of bacteria, fungi, and marine boring organisms. The borers are chiefly active below the low water level and decay is mainly above this level. Both damage in the intertidal zone. The borers chiefly belong to two groups, the Mollusca and the Crustacea. The molluscs are represented by five genera

of pholads namely *Pholas*, *Martesia*, *Xylophaga*, *Barnea* and *Lignopholas* (piddocks) and by 12 genera of shipworms of the family Teredinidae. The crustacean wood borers are mainly confined to the order Isopoda and are represented by two well-known genera *Sphaeroma* (pill bugs) and *Limnoria* (gribbles). Thirty-four species of shipworms, 7 species of piddocks, 5 species and 1 variety of pill bugs and not less than 9 species of gribbles have so far been reported from the coasts of India (Tables I, II, III). Thus, the attack on timber is the concerted effort of a heterogenous assemblage consisting of at least 55 different species of crustaceans and molluscs, besides the bacteria and the fungi. These are engaged in a relentless destruction of valuable timber thereby reducing its service life in the sea, in the brackish water and even in almost fresh water. An accurate assessment of the quantum of damage has never been made in this country. According to one estimate by Becker (1958), the cost of periodic replacement of fishing crafts alone as a result of the activity of marine timber destroying organisms is estimated at Rs. 25 lakhs. This does not include the damage done to the numerous water-front structures such as harbour construction, etc.; the estimation of which is not easy. Therefore, the total property damage caused by these pests each year must be enormous. Purushotham and Rao (1971) estimated that in India the fishing industry alone suffers an annual loss of about 10 million rupees owing to the ravages on wooden catamarans, boats, etc., by molluscan and crustacean borers along the coasts. The problem is thus as important as that of coastal erosion and should, therefore, be treated with all the seriousness it certainly deserves.

Of the marine wood-borers that occur and are active along our extensive coasts, the shipworms, piddocks and the pill bugs are responsible for most of the destruction. Gribbles though present, have not yet assumed any great importance since they occur only scarcely. Recent studies have shown that there is severe destruction of wood by crustacean borers in the coastal areas (Dharmaraj & Nair 1980) in the estuaries, backwaters (Nair 1965, 1965a 1966) and also in the mangrove forests (Dharmaraj & Nair 1980a).

Table I

*The nature of distribution of shipworms along the coasts of India*

| Name of Species                           | West<br>Bengal | Onssa | Andhra<br>Pradesh | Tamil<br>Nadu | Andaman<br>and<br>Nicobar<br>Islands | Laksha-<br>dweep<br>Archu-<br>pelago | Kerala | Karnataka<br>and<br>Goa | Maha-<br>rashtra | Gujarat |
|---|----------------|-------|-------------------|---------------|--------------------------------------|--------------------------------------|--------|-------------------------|------------------|---------|
| <i>Bactronophorus thoracites</i> (Gould)  | +              | +     | +                 | +             | +                                    | -                                    | -      | -                       | +                | -       |
| <i>Neoteredo reynei</i> (Bartsch)         | -              | -     | -                 | -             | +                                    | -                                    | -      | -                       | -                | -       |
| <i>Ducyathifer manni</i> (Wright)         | +              | +     | +                 | +             | +                                    | -                                    | +      | +                       | +                | +       |
| <i>Teredothyra excavata</i> (Jeffreys)    | -              | -     | -                 | +             | -                                    | +                                    | -      | -                       | -                |         |
| <i>Teredothyra matocotana</i> (Bartsch)   | -              | -     | -                 | -             | -                                    | -                                    | -      | -                       | -                | +       |
| <i>Teredothyra smuthi</i> (Bartsch)       | -              | -     | -                 | +             | -                                    | +                                    | -      | -                       | -                | +       |
| <i>Teredora palauensis</i> (Edmondson)    | -              | -     | -                 | -             | -                                    | +                                    | -      | -                       | -                | -       |
| <i>Teredora princepsae</i> (Sivickis)     | -              | +     | -                 | +             | -                                    | +                                    | +      | -                       | -                | -       |
| <i>Uperotus clavus</i> (Gmelin)           | -              | -     | +                 | +             | +                                    | +                                    | -      | -                       | -                | -       |
| <i>Uperotus rehderi</i> (Nair)            | -              | -     | -                 | +             | +                                    | +                                    | -      | -                       | -                | -       |
| <i>Teredo aegypos</i> Moll                | -              | -     | -                 | -             | -                                    | +                                    | -      | -                       | -                | -       |
| <i>Teredo clappi</i> Bartsch              | -              | -     | +                 | +             | -                                    | +                                    | +      | -                       | +                | +       |
| <i>Teredo bartschi</i> Clapp              |                |       |                   |               |                                      |                                      |        |                         |                  |         |
| ( <i>T. fragilis</i> Tate)                | -              | -     | +                 | +             | -                                    | -                                    | -      | -                       | -                | -       |
| <i>Teredo fulleri</i> Clapp               | -              | -     | -                 | +             | -                                    | +                                    | -      | -                       | -                | -       |
| <i>Teredo furcifera</i> Martens           | -              | -     | +                 | +             | +                                    | +                                    | +      | +                       | +                | +       |
| <i>Teredo somersi</i> Clapp               | -              | -     | -                 | -             | -                                    | +                                    | -      | -                       | -                | -       |
| <i>Teredo triangularis</i> Edmondson      | -              | -     | +                 | +             | -                                    | +                                    | -      | -                       | -                | -       |
| <i>Lyrodus affinis</i> (Deshayes)         | -              | -     | +                 | +             | -                                    | -                                    | -      | -                       | -                | -       |
| <i>Lyrodus massa</i> (Lamy)               | -              | -     | +                 | +             | -                                    | +                                    | +      | +                       | -                | -       |
| <i>Lyrodus pedicellatus</i> (Quaterfages) | -              | +     | +                 | +             | -                                    | +                                    | +      | +                       | +                | +       |
| <i>Nototeredo edax</i> (Hedley)           | -              | -     | +                 | +             | +                                    | -                                    | -      | -                       | +                | +       |
| <i>Nototeredo knoxi</i> (Bartsch)         | -              | +     | -                 | -             | -                                    | -                                    | -      | -                       | -                | -       |
| <i>Spathoteredo obtusa</i> (Sivickis)     | -              | -     | -                 | -             | -                                    | -                                    | +      | -                       | -                | -       |
| <i>Nausitora dunlopei</i> Wright          | +              | +     | +                 | +             | +                                    | -                                    | +      | -                       | -                | -       |
| <i>Nausitora fusticula</i> (Jeffreys)     | -              | +     | -                 | -             | -                                    | -                                    | -      | -                       | -                | -       |
| <i>Nausitora hedleyi</i> Schepman         | -              | +     | -                 | +             | +                                    | -                                    | +      | +                       | +                | +       |
| <i>Nausitora oahuensis</i> (Edmondson)    | -              | -     | -                 | -             | -                                    | -                                    | +      | -                       | -                | -       |
| <i>Bankia bipalmulata</i> (Lamarck)       | -              | -     | -                 | +             | +                                    | -                                    | -      | -                       | -                | -       |
| <i>Bankia bipennata</i> (Turton)          | -              | -     | +                 | +             | -                                    | -                                    | -      | -                       | -                |         |
| <i>Bankia campanellata</i> Moll and Roch  | +              | +     | +                 | +             | +                                    | +                                    | +      | +                       | +                | +       |
| <i>Bankia carinata</i> (Gray)             | +              | +     | +                 | +             | +                                    | +                                    | +      | +                       | +                | +       |
| <i>Bankia fimbriatula</i> Moll and Roch   | -              | +     | -                 | -             | -                                    | -                                    | -      | -                       | -                | -       |
| <i>Bankia nordi</i> Moll                  | +              | -     | -                 | +             | -                                    | -                                    | -      | -                       | +                | -       |
| <i>Bankia rochi</i> Moll                  | +              | +     | +                 | -             | -                                    | -                                    | +      | +                       | +                | -       |

+, Present, -, not recorded

**Table II**  
*Distribution of woodboring pholads along the coasts of India*

| Name of Species                                | West<br>Bengal | Onssa | Andhra<br>Pradesh | Tamil<br>Nadu | Andaman<br>and<br>Nicobar<br>Islands | Laksha-<br>dweep<br>Archipelago | Kerala | Karnataka<br>and<br>Goa | Maha-<br>rashtra | Gujarat |
|--|----------------|-------|-------------------|---------------|--------------------------------------|---------------------------------|--------|-------------------------|------------------|---------|
| <i>Martesia striata</i> (Linnaeus)             | +              | +     | +                 | +             | +                                    | +                               | +      | +                       | +                | +       |
| <i>Martesia fragilis</i> Verrill and Bush      | -              | -     | -                 | +             | +                                    | +                               | +      | -                       | -                | -       |
| <i>Martesia</i> sp. ( <i>obtecta</i> Sowerby?) | -              | -     | -                 | +             | -                                    | -                               | +      | +                       | -                | -       |
| <i>Pholas chuloensis</i> Molina                | -              | -     | -                 | -             | -                                    | +                               | -      | -                       | -                | -       |
| <i>Barnea birmanica</i> Philippi               | -              | -     | -                 | +             | -                                    | -                               | -      | -                       | -                | -       |
| <i>Barnea manillensis</i> Philippi             | -              | -     | -                 | -             | -                                    | -                               | +      | -                       | -                | -       |
| <i>Xylophaga</i> sp                            | -              | +     | -                 | -             | -                                    | -                               | -      | -                       | -                | -       |
| <i>Lignopholas</i> n. sp                       | -              | -     | -                 | -             | -                                    | -                               | +      | -                       | -                | -       |

+, present, -, not recorded

**Table III**  
*Distribution of crustacean woodborers along the coasts of the various maritime states in India*

| Name of Species   | West<br>Bengal | Onssa | Andhra<br>Pradesh | Tamil<br>Nadu | Andaman<br>and<br>Nicobar<br>Islands | Laksha-<br>dweep<br>Archipelago | Kerala | Karnataka<br>and<br>Goa | Maha-<br>rashtra | Gujarat |
|---|----------------|-------|-------------------|---------------|--------------------------------------|---------------------------------|--------|-------------------------|------------------|---------|
| <i>Shaeroma terebrans</i> Bate                              | -              | -     | +                 | +             | -                                    | +                               | +      | +                       | +                | +       |
| <i>Sphaeroma annandalei</i> Stebbing                        | -              | -     | +                 | +             | -                                    | -                               | +      | +                       | +                | +       |
| <i>Sphaeroma annandalei</i><br><i>travancorensis</i> Pillai | -              | -     | -                 | -             | -                                    | -                               | +      | +                       | -                | -       |
| <i>Sphaeroma tuberculatum</i> George                        | -              | -     | -                 | +             | -                                    | -                               | -      | -                       | -                | -       |
| <i>Sphaeroma walkeri</i> Stebbing                           | -              | -     | +                 | +             | -                                    | +                               | +      | -                       | +                | -       |
| <i>Sphaeroma triste</i> Heller                              | -              | -     | -                 | +             | +                                    | -                               | -      | -                       | -                | -       |
| <i>Limnoria indica</i> Becker and Kampf                     | -              | -     | -                 | +             | +                                    | -                               | -      | -                       | +                | -       |
| <i>Limnoria septima</i> Barnard                             | -              | -     | -                 | -             | +                                    | -                               | -      | -                       | -                | -       |
| <i>Limnoria tripunctata</i> Menzies                         | -              | -     | -                 | +             | -                                    | -                               | -      | -                       | -                | -       |
| <i>Limnoria bombayensis</i> Pillai                          | -              | -     | -                 | -             | -                                    | +                               | -      | -                       | +                | -       |
| <i>Limnoria pfefferi</i> Stebbing                           | -              | -     | -                 | +             | +                                    | +                               | -      | -                       | -                | -       |
| <i>Limnoria unicornis</i> Menzies                           | -              | -     | -                 | -             | +                                    | -                               | -      | -                       | -                | -       |
| <i>Limnoria platycauda</i> Menzies                          | -              | -     | -                 | -             | +                                    | -                               | -      | -                       | -                | -       |
| <i>Limnoria andamanesis</i><br>Rao and Ganapathi            | -              | -     | -                 | -             | +                                    | -                               | -      | -                       | -                | -       |
| <i>Limnoria insulae</i> Menzies                             | -              | -     | -                 | -             | +                                    | -                               | -      | -                       | -                | -       |

+, present, -, not recorded

## THE SHIPWORMS

A ubiquitous pest of all types of timber in the sea, teredo, the shipworm, causes damage worth millions of rupees every year all over the world. Hidden protectively within the 'heart' of both fixed and floating timber, and hardly visible from the outside, these borers work silently and reduce to saw dust even the most resistant timber. Rasping with their shells mechanically, these living drills draw a major part of their nourishment from the hard cellulose. Known to Pliny, Ovid and Aristophanes, shipworms are mentioned even by Homer. The early Greeks and Romans were aware of the shipworms and in fact it was Ovid in 20 BC who referred to them in his writings as *Teredin navis*. It was from this name that Linnaeus undoubtedly coined the scientific name *Teredo navalis*. The accounts of the voyages of Dampier, Cook and Drake reveal that these early navigators dreaded shipworms. Columbus lost all the ships of his fourth voyage as a result of their ravages. Thus, unaware of the danger that lurked beneath them, ancient mariners were shipwrecked in mid-ocean through the rapacity of these wood borers. Even the safety of a nation was threatened owing to the ravages of the shipworms on the wooden dykes of Holland. Despite the care and constant surveillance of harbour engineers, teredo successfully invaded San Francisco Bay in about 1921. Unseen even by experts, this exotic menace converted solid pillars and piers into weak and fragile 'honeycombs'. Along the entire seafront, bridges collapsed, piers crashed and boat hulls and wharf piling crumbled. Like an unseen typhoon it swept up the coast leaving a trail of destruction along its path. The first few waves of attack cost the United States several million dollars. In a second serious outbreak in the same locality, property worth 21 million dollars was destroyed. In 1932, piling of Canary wharves in Br. Columbia reportedly suffered a loss of \$10,00,000 worth of timber. According to estimates by the US Navy, the damage to boats, barges and bulkheads and other marine structures by borers in the US exceeds 50 million dollars every year. Naturally the destructive habits and biology of these molluscs have been the subject of much scientific interest and popular concern. Though present in all the seas, shipworms are particularly destructive in the warm, tropical waters where they eat indiscriminately every material of plant origin.

India has a long seaboard of about 8000 km taking into account the islands in the Arabian sea and the Bay of Bengal. Large quantities of timber are used for different types of water front structures such as jetties and piles, country log-rafts, like catamarans coastal and fishing vessels.



There are also harbour works, boat building and other installations of the navy and several types of aquacultural equipments all along the sea front. The money spent on all these purposes including the losses involved due to damage by microbial deterioration as well as by marine borers is enormous. The fishing industry alone which depends to a large extent on wooden catamarans and boats reportedly suffers an annual loss of about Rs. 10 millions.

Again it is not merely the monetary loss that is important, but the continuous and rapid drainage of timber from the forests particularly in countries where the supply of timber is far below the current demands. The magnitude of the problem can be realised from the fact that even the advanced countries with all the facilities they possess, such as enormous resources in the basic material, timber, technical man power and sophisticated equipments and laboratories and concentrated work on the same for over 200 years, have still not been able to penetrate beyond the fringe of the problem and these with the natural advantage they possess, namely situated as they are in a temperate region where the activity of the borers, their number and species involved are all limited. Therefore, the task that is up against the countries in the tropics such as ours, with limited resources in finance, limited area under forest and greater destructive power of borers, is very formidable indeed (Purushotham & Rao 1970).

After World War I, the magnitude of the problem drew the attention of specialists representing biologists, chemists, engineers and harbour authorities for a proper assessment of the situation, to achieve a synthesis of the existing information in this area and to evaluate the prospects of future progress. This brought about the formation of special committees to study the problem. Thus, the San Francisco Bay Marine Piling Committee, the Sydney Harbour Trust, the W F Clapp Laboratories, wood preservation units of several other countries came into being and the work by these are specially noteworthy in this connection. Similarly, the research programme of the Navy and the efforts of enthusiasts in widely separated areas and the institution of a series of test board exposure programmes gave a fillip to these studies resulting in a voluminous output of scientific literature on the subject.

Among the more notable workers in this field after the turn of the century are William F Clapp, Paul Bartsch, Miller, Kofoed and Turner (USA), Edmondson (Hawaii), Sivickis (Philippines), Iredale (Australia), Moll and Roch (Germany), Nair (India), Tchang Si, Tsi Chung-Yen, Li

Kie-min (China), Taki Habe, Kuronuma (Japan), Rjabtschikoff (USSR) and Quayle (Canada).

Excepting perhaps for iron and steel all other objects are indiscriminately attacked by marine borers. Naturally the question is asked why not replace water front structures by such materials as steel and concrete ? But with iron and steel the chief defects are their susceptibility to quick corrosion and their magnetic property which prevent their use in the construction of mine sweepers. According to Purushotham and Rao (1970) plastics and fibre glass offer as good materials for construction of light boats and life boats but they have very limited use and are restricted to ply only close to the coast or mainship. Their chief defect is that they cannot stand stormy weather and get seriously damaged in accidental knocks against hard objects like rock. Therefore, the overall picture shows that timber can be considered as an ideal material for all sea water structures. It is light in weight with high strength properties, is non-corrosive to salt water, is easy to work with and can be bent or shaped to any desired form or extended to any length by suitable joining and is also non magnetic with the result that it is used almost exclusively for the construction of mine sweepers. In most cases, it can be readily put to use in the natural form it exists as piles, bulkheads, wharves, fenders, masts, catamarans, canoes, dinghies etc. Thus, it is clear that timber is an ideal constructional material for marine water front structures and crafts.

Thus among the many questions of keen scientific interest that are encompassed by the broad field of marine biology, none has greater economic significance than the problem of marine biological deterioration. The damage caused by marine wood boring animals should have been known to man ever since his first primitive log raft was launched into the estuary or the sea. These animals have always hampered maritime activities. The ubiquitous shipworms, especially occupy a position of pre-eminence among the agents of wood destruction in salt water. Though much is known about their degradations, we are far from understanding them completely and still further from our goal of adequate and reliable control of these in all the seas and under all conditions. The problem is complicated because no two situations are identical either in the material (timber species) used or in the nature of attack (borer species) or in the manner in which the destruction takes place.

Historically, the approach to the problem of controlling these organisms has tended to be largely empirical, a trial and error evaluation

of chemical agents, techniques and structural materials to determine their efficiency in protecting marine structures against biological deterioration. The very fact that despite nearly three hundred years of serious studies on this subject in different parts of the world, only the fringe of the problem has been touched and new factors and organisms are being brought to light shows how great is the task that faces any developing country which takes up this investigation for solving its own problems.

The great amount of information which exists for such a vast field as the marine wood borers becomes so hidden in the literature and sometimes so abstruse to the non specialist that it must be brought together and suitably interpreted. Only then would it be possible to achieve an effective synthesis of the present knowledge in the field and to evaluate the prospects of future progress. Furthermore, in compounding existing information on the various aspects of their biology and ecology it would be possible to provide excellent bases from which future research can formulate control methods. There is also the need of inviting the attention of the basic researches to applied problems involved in protecting marine structures against the ravages of wood-borers. There is now the imperative need for a group of specialists representing the broad spectrum of the biological, chemical, physical and engineering sciences to come together for a proper assessment of the situation with particular reference to the ecology, biology and fundamental behaviour of these animals. It is hoped that a work of this sort will lead to a fuller and more sympathetic understanding of the problems, and interest of each group represented, with the result a common set of interest in this field will be firmly established. The preventive and control measures now in use require still further study to establish their effectiveness in a wide variety of environmental situations. Only through a more thorough knowledge of the basic biology of these organisms can we hope for ultimate understanding and control of this complex problem. A knowledge of the taxonomy, functional morphology, habits, distribution, biology and ecology of the teredines, therefore, becomes an essential pre-requisite for all further enquiries to build that frame work of understanding upon which the control of these pests ultimately depends.

The taxonomy of the shipworms was in a state of utter confusion. The need for a comprehensive treatise on the systematics and anatomy has been a long-felt need. For all concerned with the problem of biodeterioration in the sea the lack of a reliable and upto-date reference work for identification of the shipworms was a serious obstacle which

made it difficult to conduct critical field and laboratory experiments with these organisms. The monumental work of Turner (1966) entitled 'Illustrated Catalogue of the Teredinidae' has clarified the systematics of this most difficult group on the basis of careful and detailed studies of not only the pallets and shells but also on anatomical and other characters. She listed all the available names, commenting upon their synonymy and identity, summarised information on ecological and geographical distribution, proposed plausible lines of evolution among the teredinids and thus, more or less 'set the house in order'.

### PROBLEMS IN IDENTIFICATION

A perusal of the literature on teredinids will show that the taxonomy of this group has been in a state of utter confusion; the description of one species could include several allied forms. The variations in taxonomic characters exhibited by individuals are so wide that exact determination of a species may be extremely difficult. No other group seems to have a more unsatisfactory classification than the Teredinidae as pointed out by several earlier workers. The reason for this state of affairs has been (1) many of the species included under this group have been created on the basis of fragmentary material, regardless of the wide range of variation exhibited by these bivalves, (2) the locality of the type species has not been accurately determined, (3) many new species have been described on the basis of zoo-geographic provinces, without taking into serious consideration their means of dispersal; (4) authors had described several new species without reference to earlier publications which were scattered and often unavailable. This has unfortunately resulted in the creation of many invalid species.

While the taxonomy of the Teredinidae was in this state of utter confusion, Turner (1966) undertook the compilation of a comprehensive work to 'make available a catalogue of all the names used in the family Teredinidae, to illustrate as many of the type specimens as possible, giving descriptive notes concerning them, and to indicate synonyms whenever this could be done'.

Actually, shipworms have sufficient characters upon which a classification can be based, and if a large series of well preserved specimens had been available to early works in this field, much of the confusion probably would have been avoided. This was not the case, however, and many species were described on the basis of shells only or

upon a few dried specimens or on a single specimen and sometimes on a fragment of a pallet. In addition, the specimens were often taken from drift logs or from ships that had sailed in distant waters so that the origin of the specimens was unknown or the locality in error. The fact that teredinids are readily distributed by floating woods or ships was not fully realised until fairly recently and consequently many new species were described on the basis of zoo-geographic provinces (Turner 1966).

Turner (1966) had the rare opportunity, not available to many earlier taxonomists of this group of examining many type specimens, an essential pre-requisite for an attempt of this kind. Undoubtedly this work represents a milestone in the literature on the subject and will be an indispensable work of reference for all future workers.

Turner's classification differs from that of earlier workers in that she has taken into consideration some features of the anatomy of the soft parts, and also the structure and manner of growth of the pallets besides the conventional criteria for classification.

Recognising as many as fourteen genera Turner discarded the usage of subgenera owing to the occurrence of transitional species between them. Turner divided the family Teredinidae into three subfamilies, namely Kuphinae Tryon including the mud-boring genus *Kuphus* Guettard, Teredininae Rafinesque, which includes nine genera of shipworms *Bactronophorus* Tapparone-Canefrei, *Neoterodo* Bartsch, *Dicyathifer* Iredale, *Teredothyra* Bartsch *Teredora* Bartsch, *Uperotus* Guettard, *Psiloteredo* Bartsch, *Teredo* Linnaeus and *Lyrodus* Gould and the new subfamily Bankiinae which includes four genera—*Nototerodo* Bartsch, *Spathoterodo* Moll, *Nausitora* Wright and *Bankia* Gray. According to this new system, the total number of valid species in the world has been reduced to 66. This is bound to be of considerable help to all teredine workers and will enable even relatively inexperienced persons to determine the forms before them with a fair amount of accuracy. Also this revision has brought together a great deal of scattered earlier literature in the form of illustrations and descriptions of the several species. Turner has synonymised several species. There has been undue splitting of species in this group because of incorrect identification owing to the non-availability of representative series of well preserved specimens for accurate specific determination. With this new classification as the basis, the species that occur in the Indo-West Pacific have been determined.

The family Teredinidae includes the well known shipworms which are highly specialised bivalves. Unlike typical lamellibranchs, the shipworms have a naked, long, slender and cylindrical body with greatly modified but remarkably small shell-valves adapted for boring into wood. Their nearest relatives are the piddocks belonging to the family Pholadidae. Teredinidae and Pholadidae together constitute the sub-order Pholadinae of the eulamellibranch order Myoida. The sub-order Pholadina is characterised by a nearly closed mantle, a somewhat discoidal foot, reduced hinge and internal ligament, considerably small ventral and anterior adductor muscles, relatively large and powerful posterior adductor muscle and greatly modified and specialised shell valves armed with denticulated ridges over the anterior outer face and having a conspicuous pedal gape for the protrusion of the foot. Conspicuous dorsal and ventral condyles are present to facilitate rocking movements of the valves in the Teredinidae but the Pholadidae lack ventral condyles. For the insertion of the pedal muscles styloid apophyses are present beneath the umbos. While the shell valves of the piddocks can protectively cover the soft parts when retracted, those of the shipworms are greatly reduced and have thus lost their protective function but serve as effective cutting tools used for the specific purpose of excavation of the borrow. The worm-like body of shipworms extend far beyond the posterior margin of the shell, the wood into which it bores affording protection for its bare body. Additional protection is ensured by a calcareous tubing around the animal secreted by the mantle of this mollusc. While the shell valves of piddocks are provided with accessory plates, the shipworms have unique structures known as pallets located at the base of the siphons to close the burrow when the siphons are withdrawn.

### THE NATURE AND EXTENT OF DAMAGE

After settlement on wood, the shipworm larva transforms into the boring mollusc and grows with remarkable rapidity. Since the rate of growth is proportional to the destruction of timber, each shipworm during its life time destroys a column of wood of the same dimension as its largest size. The bore-hole of the piddock is much smaller than that of shipworm, yet its rapid rate of growth, and its ability to penetrate deeper and deeper in each generation cause speedy destruction of timber. *M. striata* has the ability to live in waters of very low salinity and hence this species is very widely distributed in the estuaries and backwaters of low salinity all along the coasts. On account of their density of attack, quick development, rapid

succession of generations and great tolerance to low salinities these pholads are of special importance from the point of view of timber destruction along the Indian coasts. The pill bug (*Sphaeroma*) excavates cylindrical burrows twice as long as its body and are oriented at right angles to the surface of the wood. Their dense settlement, gregarious habits and rapid rate of reproduction contribute to deeper and deeper penetration of timber. Unlike the molluscs, attack on fresh surface is effected by migrating juveniles or adults. Attack is heaviest in the intertidal zone, the maximum intensity being at half tide level. Sphaeromatids constitute a very serious threat to all available pieces of timber in the coastal areas especially the estuaries. *Limnoria* though considerably much smaller than *Sphaeroma* is capable of effecting a progressive tunnelling action on wood and make a burrow several times the length of its body. Innumerable small holes, produced by these give the wood a sponge-like texture and lace-like appearance. Burrows usually follow the grain and are close to the surface of the wood.

Thus, the nature of damage by the molluscs and crustaceans is different, producing different effects on timber. This enables them to share effectively, without serious competition, this common substratum which is limited in extent. The crustaceans work from the outside and the molluscs, particularly the shipworms penetrate deep into the 'heart' of timber. The combined action of these two groups of borers converts the wood into a highly porous, weak and fragile mass. The crustaceans have the added ability to enter even creosoted shell of treated timber which the shipworm larvae are unable to do.

Besides the woodboring animals, wood-infesting bacteria and fungi, especially in the Ascomycetes and in the Deuteromycetes (Fungi imperfecti), actively participate in a sort of 'conditioning' of the timber, preparing it for the subsequent attack by borers. This is a biological phenomenon of considerable importance in the ecology of marine borers. The activity of the fungi leads to a kind of deterioration called 'soft rot' which is a sort of superficial softening of the wood through a cellulolytic process. These fungi which are resistant to preservatives, release a strong cellulase which hydrolyses the unlignified cell elements leading to the softening and disintegration of the outer tissue of timber. Even though the magnitude of damage by these is not spectacular and may not even be noticed by the layman, the importance lies in the fact that their silent and steady activity soon after submergence of timber prepares it for attack by the crustaceans and molluscs. The borers in turn help the fungi to spread

deeper and deeper into the timber, thereby enabling them to expand the field of operation from superficial layers to its very core. The relation between gribbles and these fungi seems to be of a symbiotic type. The fungal infestation on light timbers of the catamarans, dugout canoes and other fishing crafts, according to Becker and Kohlmeyer (1958) is not the usual superficial type, the penetration being deep, affecting the entire log. The periodic drying of these logs accelerates the spread of the fungal hyphae which get effective ventilation through the large vessels of these light timbers. Preliminary studies have revealed the existence of several species of marine fungi in test panels.

### THE PATTERN OF DISTRIBUTION OF TIMBER BORERS ALONG THE COASTS OF INDIA

The pattern of distribution of the teredine, pholadid and crustacean wood-borers along the coasts of India and in the Indian ocean islands is shown in tables 1-3 respectively. This list is not a complete representation of the species that occur since virtually nothing is known regarding their incidence from many areas along our coasts. This list is compiled from observations in certain selected localities or from the vicinity of research institutions which carry out studies on marine boring animals. Reports from other localities are mostly based on a few collections which are not expected to yield satisfactory results. Nevertheless, the members noted in the list would represent the most dominant forms, and therefore the highly destructive species in the respective localities. Some species such as *Dicyathifer manni*, *Teredo furcifera*, *Lyrodus pedicellatus*, *B. carinata*, *B. rochi* and *Martesia striata* are widely distributed along our coasts. Except *Neoterodo reynei*, *Bankia rochi*, *B. fimbriatula*, *Nausitora fusticula*, *Nototerodo knoxi*, *Spathoterado obtusa*, *Teredothyra matocotana*, *Teredora palauensis*, *Teredo somersi*, *Teredo aegypos* and *Nausitora oahuensis* all others occur along the Tamil Nadu coast. The next higher species representation is along the coast of Andhra Pradesh from where as many as 18 species of molluscan wood-borers have been reported. This is roughly proportional to the amount of work carried out from the coasts of the various states. The number of species occurring along the coasts of other states might also be much more than that indicated and it necessitates further detailed surveys along the unexplored coastal zones.

Some interesting features could be observed from the record of occurrence in various localities and from the nature of distribution of the



wood boring pests along the coasts. Many species show a discontinuous distribution, e.g. species of *Teredothyra*, *Uperotus* etc. *T. smithi* has been collected from Tamil Nadu, Gujarat and the Lakshadweep coasts only. Species of *Uperotus* are characteristic in their occurrence only along the coasts of Andhra Pradesh, Tamil Nadu and in the Lakshadweep Archipelago. *T. furcifera*, one of the widely distributed species has not so far been reported from West Bengal and Orissa. *Teredo fulleri* seems to be a characteristic borer occurring in the Gulf of Mannar and the Palk Bay coasts of Mandapam, neighbouring islands and in the Lakshadweep Archipelago. *Teredo bartschi*, *L. affinis*, *Nototeredo knoxi*, *Nausitora dunlopei*, *N. fusticula*, *Bankia fimbriatula*, *B. bipennata*, *Xylophaga* sp and the *Barnea birmanica* also have been reported from the east coast of India but not so far from the west coast.

Table II presents the nature of distribution of crustacean borers. Throughout the west coast and along the east coast bordering Andhra Pradesh and Tamil Nadu, *Sphaeroma terebrans* has been reported to occur abundantly in brackish water and other estuarine localities. *S. annandalei* var. *travancorensis* occurs along the Kerala and Karnataka coasts and *S. tuberculatum* is known only from its type locally, Tuticorin. *S. triste* occurs along the coast of Tamil Nadu and in the Andaman and Nicobar islands. *S. walkeri* has been collected from Maharashtra, Kerala, Tamil Nadu and Andhra coasts. Of the nine species of limnorids, all but *Limnoria tripunctata* and *L. bombayensis* occur in the Andaman Nicobar Islands. Limnorids are apparently not very active along the coasts of Kerala, Karnataka, Andhra Pradesh and Gujarat. In some localities such as Bombay, Madras, Mandapam and the Lakshadweep Archipelago the incidence of limnorids has been steadily increasing and this is likely to cause damage to the superficial layers of exposed wood. *Limnoria indica* and *L. bombayensis* are the destructive species along the coasts of the mainland.

Preliminary studies in the Pichavaram mangrove forests on the south-east coast of India have shown that wood-boring animals are very active there and these represent a perennial source for infestation along the coasts. At least six species of shipworms namely *Bactronophorus thoracites*, *Lyrodus pedicellatus*, *Teredo furcifera*, *Nausitora hedleyi*, *B. carinata* and *B. campanellata*; the pholads, *Martesia striata*, *Barnea birmanica*; and the pill bugs *Sphaeroma terebrans* and *S. annandalei* have so far been recorded.

Studies on the nature of infestation in the aquafarms along the southeast coast revealed that several species occur there and rapidly reduce the service life of timber used. The different types of equipments used in these operations experience different types of borer activity depending on the level and locality of exposure. In the Karapad creek oyster farms at Tuticorin the most dominant and destructive species are *Teredo furcifera*, *Lyrodus pedicellatus*, *L. affinis* and *Martesia striata*. The teredines infest the piles from surface to mudline with dense settlement near the bottom and the attack of pholad is dense a little below the low water mark.

In the pearl oyster and seaweed farms located along the coast of Krusadai Island in the Gulf of Mannar the relative abundance of different species of shipworms is as follows: *Teredo furcifera*, *Teredo fulleri*, *Lyrodus pedicellatus*, *Teredo triangularis* and *T. bartschi*. The most affected are the sea weed farms where numerous wooden piles have been installed for the rope culture of *Gracilaria* spp. and other algae. In these farms, gibbles (*Limnoria* spp) also are active. In the oyster and fish farms wood is riddled mostly by *Teredo furcifera* and *Lyrodus pedicellatus*. Stakes used here are serviceable only for a brief period of about 4 months despite the traditional protective coatings applied on them prior to installation. Along the coast of Hare Island *T. furcifera*, *T. bartschi*, *T. fulleri* and *L. pedicellatus* are destructive. Stray occurrence of *Limnoria* has also been noticed.

In the aquafarms in the Vellar Estuary near Porto-Novo occur *L. pedicellatus*, *N. hedleyi*, *S. terebrans* and *S. annandalei*.

The Nethravathy-Gurupur estuary and the coastal waters of Mangalore on the south-west coast, harbour *L. pedicellatus*, *Dicyathifer manni*, *N. hedleyi*, *M. striata*, *S. terebrans* and *S. annandalei*. Most of these species are capable of tolerating the fluctuating salinity encountered in such situations.

The Vembanad and the Ashtamudi lakes on the Kerala coast with extensive fishery and aquacultural possibilities contain such species as *L. pedicellatus*, *T. furcifera*, *Nausitora hedleyi*, *Dicyathifer manni*, *S. terebrans*, *S. annandalei* and *M. striata*. *B. carinata* is seen occasionally near the bar mouth. *N. hedleyi* is most destructive in the Vembanad Lake. In the Vizhinjam Bay on the south-west coast the wooden rafts used for the culture of the pearl oysters and mussels are attacked by *L. pedicellatus*, *B. carinata*, *B. campanellata*, *T. furcifera*, *L. massa* and *M. striata*.

Timber structures and drift wood in the coastal waters of the Lakshadweep Archipelago were found to harbour not less than 19 species of molluscan borers. The dominant ones in the region are *Teredo fulleri*, *T. clappi* and *Lyrodus massa* (Nair & Dharmaraj 1983).

## NATURE OF INFESTATION AND VERTICAL DISTRIBUTION

Information on the nature of vertical distribution of woodboring animals is of considerable value since the degree of deterioration at different levels along a pile is based on the intensity of settlement and growth at these levels. Further, a study of the varied biological relation which permit a heterogenous group of boring animals to share a common and limited habitat will be of interest ecologically.

Depending on the depth of exposure of the wooden equipments, the intensity of destruction would vary. Piles and stakes are infested from high water mark to the very mud-line since they run throughout the water column and are subjected to settlement by all kinds of borers. Shipworms generally infest the wood from surface to bottom with dense settlement a little above the mud-line both in coastal waters and in estuaries (Nair 1966, Nair & Dharmaraj 1979). In the estuarine shipworm *N. hedleyi* significant differences have, however, been noticed in the intensity of settlement at the different levels during the different periods of the year (Saraswathy & Nair 1969). During the monsoon period distinctly greater numbers settled over the bottom. During the post-monsoon the nature of settlement was different with greater strike sub-tidally. This shift in the nature of settlement is probably due to the influence of some ecological factor most probably salinity. It was not evident from tests whether the larvae were distributed equally throughout the water column and only those at certain levels were successful in settling and surviving giving the impression of greater settlement at those levels. This case suggests that borer activity though generally considered most abundant near the bottom need not necessarily be so at all locations. *Teredo furcifer* has been found to settle abundantly near the low water mark in Visakhapatnam Harbour (Nagabhushanam 1960). *Martesia striata* is active throughout the water column in coastal waters, but the region of dense attack in estuarine environments is near the bottom (Nair 1986, Dharmaraj & Nair 1979). Thus in estuarine areas, attack by both shipworms and piddocks is concentrated towards the bottom and this region of the piles becomes weak and breaks off. Sphaeromatids and limnorids abound the timber structurers at the intertidal zone. In the backwater and estuarine

environments piles and pillars subjected to sphaeromatid attack assume characteristic hour-glass-shape in course of time on account of the combined action of borers and the mechanical action of waves. Floating rafts in coastal areas are likely to be invaded by piddocks, limnorids, and shipworms in the areas of contact with water but the intensity of destruction would be less severe than in fixed structures. Rafts exposed in brackish water and estuarine areas are susceptible to heavy attack by sphaeromatids. Cages and rafts used in mariculture operations would experience varied types of borer activity depending on the depth of exposure and the species of borer present in the area.

Woodboring sphaeromatids are active in brackish water but the allied ispod *Limnoria* is quite different in its preferences, none being recorded as destructive in low saline areas. *Martesia striata* is able to thrive in a wide range of salinity but *M. fragilis* occurs only in the offshore waters where conditions are quite constant. Of the dominant species of teredinids in India *L. pedicellatus* has been highly destructive in estuaries, mangrove swamps, backwaters and also in coastal waters. *T. furciferta* is active mostly in coastal waters and occasionally intrudes into estuaries also when conditions are favourable. *B. carinata* and *B. campanellata* are destructive in coastal and offshore waters. Species of the genus *Nausitora* have mostly been collected from estuaries, backwaters and mangroves. *B. thoracites* and *Dicyathifer manni* are also normally confined to brackish water localities. *T. clappi*, *T. bartschi*, *T. fulleri*, *Lyrodus massa*, *B. rochi*, *Teredora princesae* and *Uperotus rehderi* are also destructive in certain coastal and offshore localities.

### SEASON OF SETTLEMENT

The time of settlement is of special interest apart from its biological importance because it is then that the infestive free-swimming larvae come into contact with fresh surfaces and experience the effect of preservatives used on them. Precise knowledge of the times of settlement of the different species in a locality is of importance in connection with such operations as replacements, dry docking, and repainting of wooden boats, pile driving etc. Biologically, the period and extent of settlement are significant since they are reflections of the breeding season and a reliable measurement of the breeding success. This is due to the fact that there is the possibility of spawning without settlement. In certain instances detailed studies on the season of settlement of timber boring organisms have shown effective ecological adjustments, the different species

occurring in an area showing interrelationships so that interspecific competition is reduced to a minimum through characteristic zonation in settlement (Nair 1959). Alternation of breeding periods prevents simultaneous settlement, the different species occurring in the area settling over the limited amount of available timber at different times of the year leading to a succession in settlement (Nair 1965). The breeding activities of closely allied species show difference and even those of the same may vary according to the hydrographic conditions prevailing in the area (Nair 1965). The density of distribution of shipworms has fluctuated over long periods and within the same period their attacks have differed considerably in various locations along the same stretch of coastline (Becker 1958).

This aspect of the biology of shipworm has been the subject of study in three localities along the coast of India. At Cochin Harbour, Nair (1965) observed that *T. furcifera* settles chiefly during the hot, highly saline pre-monsoon period February to June, with sparse settlement during the early part of the monsoon and later part of the post monsoon periods. The settlement of the estuarine species *N. hedleyi* is confined to the low saline periods of the monsoon and the post-monsoon periods (August-February) with apparently no settlement during the premonsoon. Thus the settlement of these two species alternates in Cochin Harbour. Nagabhushanam (1962) noted at Visakhapatnam (east coast of India) the occurrence of *T. furcifera* on test panels throughout the year with a maximum attack during the summer months between March and June with a peak in May of 1956. The difference noticeable in the nature of settlement along the south-west coast and east coast of India may be explained on the basis of the prevailing hydrographic conditions.

The period of settlement of *B. campanellata* common at Visakhapatnam is from August till February and the species is absent from the panels during March to July, the maximum intensity of attack was noticed during November to January. Nagabhushanam (1962) pointed out a direct relationship between the relative abundance of shipworm settlement and temperature and salinity. The comparatively smaller attack rate by *Teredo* during the winter months (November-January) was attributed to a biological competition with *B. campanellata* whose intensity was greatest during the winter months.

Confirmatory studies of Saraswathi and Nair (1969) on the settlement of *N. hedley* showed that it is strictly seasonal from July to

February with November representing the peak period. Based on a detailed study of the frequency of occurrence of veliger larvae in the plankton, presence of post-settled stage on test panels and the condition of the gonads of the adults, Nair (1955, 1957) concluded that the period *July-August* is the best for larval development and attachment of *B. carinata* at madras.

In general, it seems that under tropical conditions the settlement continues without interruption in the marine zone (Nair 1957). Even under such conditions the extent of settlement varies from month to month. In special habitats such as estuaries, the influence of salinity may prevent the uninterrupted breeding and settlement of some species (Nair 1965). The number of generation which would be produced in a single season varies with the time required for the species to grow to maturity and with the length of the period during which suitable conditions persist. In warmer regions development is more rapid and many generations may be produced each year. In regions other than the tropics temperature appears to be the principal conditions determining periods of breeding. Adult animals can frequently survive under extremes of temperature which are unfavourable for reproduction. Consequently a species may maintain itself where conditions are suitable for reproduction during only a small part of the year.

In making generalisations one should not forget that the variables other than temperature may be directly or indirectly influencing events, for example, especially along the south-west coast of India, rainfall reflects the passing seasons, causing great reduction of, as well as fluctuations in, salinity and so creating conditions unsuitable for breeding and settlement of certain species.

Events leading to settlement are dependent on the interplay of a large number of factors including the physiological characteristics of the species involved, their geographical distribution, local variations in the character of temperature changes and the seasonal influence of other and less obvious aspects of the environment (Nair & Saraswathy 1971).

#### DURATION OF LARVAL PERIOD

The average duration of the free-swimming stage of the veliger is apparently constant for a given species at a given locality. Naturally this period is shorter in the warmer waters of the tropical and subtropical regions. The duration of free-swimming period in the non-incubatory

species vary. In *B. campanellata* about 17 days in Madras (Nair 1956) and in *B. campanellata* about 13 days in Visakhapatnam (Nagabhushanam 1959). In the tropical incubatory species *T. furcifera* larvae attract timber normally between 24-72 hr and they need no food prior to settlement (Karande *et al.* 1968). The source of energy for the larvae has been attributed to the stored glycogen (Lane 1955).

Preparatory to settlement, they crawl over the surface of the wood searching, probing and prospecting the surface for a suitable spot, the period of crawling varying greatly, extending even upto an hour. During this period, activity diminishes till the foot comes to rest in a depression on the surface. Subsequently, the shell is lifted while the byssus attachment is made. The velum is retracted. At this stage the velum and the foot are both functional so that the larvae can alternatively swim and crawl, a stage aptly termed 'pediveliger'. If a suitable substratum is not available the free swimming larval life could be continued for several days, and the pediveliger has the ability in certain species to postpone metamorphosis (Nair 1956). This ability is probably of survival value permitting the larvae to cover a wide area in their search for the appropriate substratum. In some species such as *Lyrodus pedicellatus*, larvae denied access to wood lost their ability to penetrate it within 4 days and invariably died in 2 weeks. The infective period for this species was the first 96 hrs after release from the parent (Lane 1959).

Larvae are attracted to wood, at least in the sense that, should they chance to encounter it, they remain upon it and there metamorphose. The wood must be properly conditional for penetration and this involves among others saturation with water and the development of suitable micro-flora and fauna (Nair 1965).

### GROWTH RATES

Information on growth rates is important because growth rates are directly related to the damage done to timber. Each shipworm destroys a column of wood of the same dimension as its largest size. Comparison of growth rates of different species of shipworms recorded from different kinds of timber from different localities do not seem to have much meaning.

The growth rates of different species occurring in widely separated places are likely to be different and even in the consideration of growth rates of the same species, unless the latitude, substrate involved etc., are the same the values are likely to vary. Growth rates need not be identical

even when latitude, species and substrate are all the same during the different seasons since they are likely to be influenced by the prevailing hydrographic conditions. The same species has different rates of growth at different localities of the same estuary during the same time depending on the number of individuals which have settled. In *Bankia carinata* of Madras the growth rates recorded by Nair (1960) are as follows: 9mm in 17 days, 23mm in 32 days, 142mm in 68 days, 224mm in 95 days, 257mm in 125 days, 274mm in 165 days, 290mm in 191 days and 302mm in 219 days, representing an average rate of boring of 4.3cm/month. Growth was very rapid during the first 90 days and thereafter slackens and the trend indicates that growth becomes negligible at the end of about 220 days. Retardation in growth has been attributed to depletion of woody materials on the panel as a result of overcrowding. In *B. campanellata*, a growth of nearly 51.5mm/month has been recorded with a maximum of 60mm in December, and *T. furcifera* attains a length of 90mm at the end of 5 months in Vishakhapatnam (Nagabhushanam 1959). Studies on the estuarine shipworm *N. hedleyi* for six months indicated that growth was rapid between 45 and 153 days after settlement with the maximum growth between 105-120 days of age. Thereafter growth slackened and the trend indicated the growth was negligible at the end of 150 days (Saraswathy & Nair 1974).

### DISPERSAL OF SHIPWORMS

Notwithstanding the stationary, hidden life within the confines of their wooden burrows, shipworms are distributed far and wide through their free swimming larval stages. While some species liberate eggs into the water, others brood the eggs, the veligers being released when ready. During the free swimming period which may last from a few hours to even a month depending on the species and the region, the larvae are drifted about and widely transported in the surface currents.

Three types of larval life could be recognised among shipworms; (i) oviparous (ii) short-term larviparous (iii) long-term larviparous. The pattern of distribution of many species is based on the temperature and salinity requirements of the larvae during their plankto-trophic life. The members of the genus *Nausitora*, generally denizens of brackish water habitats, are oviparous and fertilisation is external. Since the larvae are incapable of tolerating high salinities (Saraswathi & Nair 1974), those larvae transported by the medium into areas of high salinities would perish. This accounts for their occurrence in isolated pockets of brackish



water. In this case colonisation of new areas could be effected only by the adults which are more tolerant to higher salinities than their planktonic larvae and are transported through the agency of drift wood etc. to a suitable brackish water environment within the life time of the adult. The picture of the distribution of this genus seems to support this contention. *N. hedleyi* and *N. dunlopei* have so far been reported only from isolated brackish water areas along the tropical Indo-West Pacific, and *N. dryas* and *N. excolpa* along the tropical eastern Pacific.

Species of *Bankia* are also oviparous with protracted planktotrophic larval stages but their habitat is normally marine. In this case the limiting factor in distribution is apparently temperature. Therefore, species characteristic of higher latitudes are restricted in range incapable of spreading into sub-tropical or tropical areas as may be seen in *B. setacea* of eastern and northern Pacific and *B. gouldi* of northern and western Atlantic. Similarly the tropical species *B. carinata*, *B. campanellata* and *B. bipennata*, though established all round the world are restricted by the temperature factor and have apparently not spread beyond the subtropical zone.

While discussing the dispersal of marine species with long-term larviparous young such as the members of the genus *Lyrodus*, and *T. furcifera*, *T. clappi*, *T. somersi* and *T. bartschi*, Turner (1966) enumerates the reasons for their worldwide distribution: (i) in common with other shipworms the adults can be transported to great distances in ships or floating wood; (ii) the young are protected within the parents during the early critical stages of development; (iii) the larvae are not spawned unless optimal conditions for their survival exists; (iv) being further developed the young are less sensitive when extruded; (v) the larvae are ready to settle shortly after they are extruded and not carried away from the floating log or ships from which they emerged; and (vi) most wooden ships and pieces of driftwood are covered with a good growth of hydroids, bryozoans, algae and other organisms which form a protective forest cover within which the larvae can swim until time of settlement.

Dispersal of shipworms is commonly effected by driftwood or through the agency of wooden hulls of ships, etc. Extended and intensive intercourse between nations in the long years of maritime activity contributed to the spread of this menace in widely separated places, chief agents in transport being the hulls of wooden ships, wooden sea water

tanks of ships and log-booms. Rapid increase in maritime shipping in the past two centuries helped this transport. The European species reached the American shores during World War I. Ships transported at least one species to South Africa, China, Japan and Australia. It is also true that infested driftwood carried by the surface currents played a major role in their wide distribution. Floating nuts and seeds drifting passively in the surface currents distributed at least one nut infesting species (*Uperotus clavus*) in the region between east coast of Africa and the Philippines. From the Philippines and neighbouring areas many species have reached Hawaii in drift-wood and in light ocean going craft of wooden construction. The whole Indo-West Pacific area contains many common species. There is also the possibility of the release of larvae from infested waterlogged wood lying on the sea bottom. This picture of wide distribution can be explained either on the basis of passive dispersal of free swimming larvae and of adults through the prevailing surface currents and drift-wood respectively or through the active transport of adults by ships. Thus a few species like *B. carinata*, *B. campanellata* have succeeded in establishing themselves around the world in both tropical and sub-tropical waters. It has also been noted that the larvae can be effectively transported amidst the thick growth of fouling organisms that accumulate over the bottoms of the steel hulls of ships. Larvae taken in the feeding current in one locality by certain attached species on the ships hull may pass through the alimentary canal apparently undamaged and emerged along with the faecal pellets when the ship reaches another locality. These observations reveal that several species of shipworms have been dispersed and others may be expected to be so, over wide areas of the oceans. Certain species such as *T. princesae* can spend the life cycle from larva to adult in the open sea supported by some suitable flotsam without making contact with stationary structures in coastal waters. The larvae of these 'seasoned ocean travellers' can endure long enough to contact drifting timber and continue a chain of seafaring generations. While specimens from drift logs and wooden hulls may serve as a record of the locality they may not represent established fauna. Conditions must be sufficiently suitable for spawning, for the survival of the larvae and successful penetration of the wood before a species can become an established member of any local fauna (Turner 1966).

## ECOLOGICAL FACTORS

There are several environmental factors which affect the natural populations of shipworms. These are the physico-chemical variables of the sea water, such as temperature, salinity, oxygen tension, turbidity and pollutants, the presence and intensity of fouling organisms, the nature of the wood, depending on the species of timber, its softness and orientation of the grain; the length of exposure of the wood sample in water; the presence or absence or the nature of preservatives used on it; the location of the wood in relation to tidal changes, whether or not it is periodically exposed to desiccation; the orientation of the sample in relation to depth; nature of the bottom; mechanical effects of currents, their velocity, conditions of illumination; the interaction of the species of wood boring animals present in the area; the availability of a suitable substrate during settlement, the effectiveness of local larval sources in the case of shipworms and the presence or absence of predators and parasites. The occurrence, abundance, and so the intensity of attack in any locality is dependent on these factors which vary widely from year to year. Probably there are several more, but these factors are the most important. Variations in the borer populations from year to year in any locality are no doubt due to a very involved association of these factors some of which occasionally stand out as the most responsible ones while other factors, none attaining conspicuous importance by itself, may collectively exert as much or more influence than more prominent and easily followed factors. It is by a constant shifting of importance on these factors and new alignments in their association owing to ever varying conditions that accounts for variations in local abundance of shipworms. This would explain periodically recurring devastations separated by often lengthy intervals of comparative freedom from attack. Reasons for increased attacks have not all been investigated but they may be different in different localities such as lowered salinity caused by inflow of fresh water or reduced rainfall causing an increase in salinity, high temperature and dry summers or an unusual increase in water temperature.

Temperature is a major influence on the activities and distribution of shipworms and is a limiting factor in growth, reproduction and distribution. There is a complex correlation between the biological effects of temperature and salinity, the former can modify the effects of the latter and change the salinity range of an organism. Depending on the temperature tolerance some species are characteristically restricted to the cold waters of higher latitudes, others are typically inhabitants of sub-

tropical areas while a majority are denizens of the tropical regions illustrating the general rule that tropical biotopes are typically rich in species. Thus along the coasts of Norway 3 species of shipworms occur (Nair 1959), 5 species in the Mediterranean, 7 species in Hawai and Midway Islands, 16 species in the Philippine Islands and as many as 34 species in India. Uniformly high tropical temperature hasten metabolic activities and accelerate growth rates, leading to attainment of sexual maturity at a surprisingly early age. Some species breed almost continuously and several generations are produced in rapid succession within a single year. Such speeding up of life histories favour the acceleration of the evolutionary process. This probably accounts for the richness in the represented species in the tropics.

The uniformly high temperature of the tropics can stimulate sexual activity, accelerate development of gonad, hasten maturity and shorten the free-swimming larval period. These contribute towards the production of several spawning in a single year leading to an almost continuous settlement of waves of borers which bring about speedy destruction of timber. Similarly the period of free-swimming of larvae and growth rates may also be influenced by temperature.

### SALINITY

Salinity affects the organism by influencing the density of the medium and through variations in osmotic pressure. In tropical estuaries the wet season with low salinity is in summer. Though lowered salinities may be tolerated at summer temperatures, where this drops suddenly and then continues at a low level for a long period, stenohaline species are likely to be killed.

The reaction of shipworms to different salinities varies widely. Some species can tolerate only high salinities others can tolerate a fairly wide range of salinities while a few are capable to enduring very low salinities and even freshwater. Further the salinity tolerance of the same species may vary according to the geographic location depending on the prevailing temperature and may even vary in the same locality during the different seasons of the year. This is due to the existence of a complex correlation between the biological effects of temperature and salinity, the former having the ability to modify the effects of the latter and change the salinity range of an organism (Kinne 1963).

According to Nagabhushanam (1962) *B. campanellata* in Visakhapatnam occurs in areas where the salinity is between 21 and

34‰. The magnitude of strike decreased with decreasing salinity. The genus *Nausitora* is generally confined to brackish water although some species have occasionally been taken from marine habitats (Nagabhushanam 1960, Nair 1954). In *N. hedleyi* a majority of the adults are typically euryhaline capable of enduring a wide range of salinities (0.65-33.68‰) but the breeding is apparently restricted to the low saline period.

The first record of this genus was by Wright (1854) who obtained specimens of *N. dunlopei* from freshwater (?) 150 miles above the mouth of the Ganges. Rajagopalaienger (1961) reported *N. lanceolata* (= *N. dunlopei*) from Sajnaikhalı 24 Paraganas District of W.Bengal, *N. hedleyi* occurs in the low saline waters of the Pulicat lake on the east coast (Nair 1963). Records show that this genus is sensitive to higher salinities and so restricted to estuarine areas. *N. hedleyi* can apparently withstand wider changes of salinity than other species of the genus and is capable of tolerating much lower salinities than typical marine species. Great damage can be expected from species of this genus in the low saline localities of river mouths etc. where species sensitive to low salinities cannot survive. Thus the genus *Nausitora* effectively occupies an ecological niche and so extends the zone of operations of shipworms to river mouths and even further upstream. The construction of dams and barrages for hydroelectric power and for irrigation across the rivers and the consequent check on the river flow may lead to greater spread of the tidal water upstream and this can result in the extension of shipworms activity. Observations at Cochin on *N. hedleyi* suggest that for the early development of the species the most suitable salinity range is from 11.24-14.54‰ (Saraswathy & Nair 1974a). Above and below this, segmentation is abnormal and the percentage of normal embryos decline. In salinities lower than 4.36‰ and above 27.14‰, there is no evidence of development (Nair & Saraswathy 1971). Most species of shipworms require normal marine conditions for successful spawning but the adults may withstand unfavourable conditions by closing the burrow with the pallets. They are also able to utilise the stored glycogen under anaerobic conditions.

### OCCURRENCE IN FRESHWATER

Shipworms are not limited to the sea and brackish water. *Nausitora dunlopei* has been reported in timbers in the river Comer, a tributary of the Ganges 150 miles above the mouth where the water is fresh (Wright 1864). Shipworm activity has also been reported in almost freshwater up

the Mississippi, in the Panama Canal, Chulalongkorn Lock in Thailand, in the Sacramento river California, etc.

### THE PRIMARY FILM

Contrary to the finding of Nagabhushanam (1959a), Karande *et al.* (1968) noted that the settlement of *T. furcifera* in Bombay harbour was independent of both light intensity and the presence of a primary film comprised of bacteria, algae or fungi since larvae settled and grew even on sterilised 'clean' timber blocks, the latter being maintained in sterilized seawater for a number of days. According to Karande *et al.* (1968) it is the softening effect of sea water, rather than the microfilm that helps borer larvae to abrade the wood.

Wood-infesting bacteria and fungi especially in the Ascomycetus and in the Deuteromycetus (*Fungi imperfecti*) participate in a sort of 'conditioning' of the timber preparing it for the subsequent attack by borers. This is a biological phenomenon the importance of which has only recently been stressed (Kohlmeyer 1963). The activity of these fungi leads to wood deterioration called 'soft rot'. These are resistant to preservatives and release a strong cellulase which hydrolyses the unignified cell elements leading to the softening and disintegration of the outer tissues of timber. Even though the damage is not spectacular and may not even be noticed by the laymen, their steady activity following submergence prepare timber for attack by crustacean and molluscan borers. These in turn help the fungi to spread from superficial layers to the very core of the timber. Fungi infestation on light timbers of catamarans, dug-out canoes and other fishing crafts, according to Becker and Kohlmeyer (1958) is not of the usual superficial type, the penetration being deep, affecting the entire log. The periodic drying of these logs accelerates the spread of the fungal hyphae which get effective ventilation through the large vessel of these light timbers. Our studies at Cochin harbour have revealed the existence of several species of marine fungi in the timber test panels such as *Gnomonia longirostria* Cribb & Cribb, *Halosphaeria quadricornuata* Cribb & Cribb, *Torpedospira radiata* Meyers, *Corollospora pulchella* Kohm, Schmidt and Nair *Lulworthia* sp (Nair 1968). The exact role played by these fungi in the ecology of shipworms is under investigation.

### RATE OF WATER FLOW

The rate of water flowing over them is of considerable significance in the distribution of sessile marine invertebrates. The effects of water currents

upon the rate of settling of *Teredo furcifera* and *B. campanellata* on timber have been studied by Nagabhushanam (1961). He found that these borers "require some water current velocity for settling on timber and that they settle more rapidly, when the waters are flowing than when waters are still". Probably a flow is beneficial in carrying larger number of larvae to the test site.

### TURBIDITY

Turbidity is an important factor especially for the inhabitants of shallow coastal waters and of estuarine regions where turbid conditions are often created by deforestation, heavy rains and consequent river discharge, wind action, dredging operations, boat and ship traffic, etc. In the tropical regions during the time of the monsoon highly turbid conditions may exist for weeks or even months. Presence of detritus affects organisms by providing food and influences the transparency of the water affecting the rate of organic production. The apparent immunity from borer attack of the base of the piles of Kidderpore Docks in Calcutta has been attributed to the silty and muddy bottom (Devenish-Meares 1904). Nair (1962) has reported that both crustacean and molluscan borers are much less active where the bottom sediments are being continually churned up by propellers.

### POLLUTION

The habitat of shipworms, especially those in estuaries, harbours and similar situations are subject to the effects of pollution either through industrial wastes or by human sewage. The pollutants may consist of solid matter or soluble chemicals of a toxic nature the presence of which can affect the organisms directly or they influence the water, for instance its oxygen content. Some effluents may be beneficial in small quantities. Waters with heavy sewage pollution or which are influenced by  $H_2S$  as in some of the backwaters of Kerala are comparatively free from shipworms. That pollution of harbour waters is unfavourable for shipworm activity is evident from the observations of Nair (1962) who found that the activity of mat-forming fouling organisms over timber structures in areas with sewage pollution is beneficial since these organisms form a protective cover against the settlement and penetration of borers.

## EFFECTS OF FOULING

Dense fouling accumulations over underwater surfaces can effectively inhibit the attack by both crustacean and molluscan borers. Of the different groups of fouling organisms, the barnacles are perhaps the most effective agents hindering attachment of shipworm larvae although other mat-forming organisms may also serve as protection (Nair 1962). Nagabhushanam (1960a) found that fouling has a profound effect on the attack of marine borers, fouled blocks showing only about one-ninth as much attack as did weekly cleaned panels.

## PARASITES AND ASSOCIATES

The role of molluscs as hosts of zoo-parasites has been well-known for over two centuries yet no detailed study has been made of the parasites and associates of wood boring organisms. Several organisms are associated with shipworms in their natural habitat but so far no cause and effect relationship has been established. There is always the possibility that certain parasites and predators may be utilised as effective agents of biological control. Our recent studies have brought to light an interesting fauna of protozoan associates from shipworms (Nair 1968). So far we have collected species belonging to Boveridae, Urceolaridae, Thigmophyridae, Spirostomidae, Lichnophoridae, Hysterocinetidae, Stenotoridae, and also *Zoothamnium*, *Lagenophrys* and *Trochilioides*. Detailed taxonomical and ecological studies on these protozoans are in progress. Association between hydroids of the genus *Eugymnanthea* and bivalves has been reported from time to time. The first record of this interesting association between a hydroid recognizable as belonging to *Eugymnanthea* and shipworms has been by Santhakumari and Nair (1969) who reported its occurrence in *N. hedleyi* and *T. furcifera* from Cochin backwaters. This hydroid, however, differs from all other species of *Eugymnanthea* in the nature of the polyps which are not solitary but branched. The incidence of the hydroid is seasonal in this locality during November-May with nearly 100% infestation in shipworms collected during December to March. The number of colonies present in a host ranged from 1 to 80 and the hydroid was found exclusively attached to ctenidia. Attachment is by a basal disc which is firmly implanted within the tissue of the ctenidium and is further strengthened with the aid of protrusions or 'holdfasts' that project from the basal disc—a feature very different from that in species of *Eugymnanthea* previously studied. There is no evidence of fatal damage to the host tissue. The more intimate nature



of the association and the tendency towards colony formation are interesting features. Studies of Santhakumari (1970) on the medusa of this form led her to assign this to *Eutima commensalis*. Turbellarians (polyclada), predaceous polychaetes, commensalic copepods and amphipods have also been recorded from the burrows of shipworms.

With their characteristic boring habits and a boundless appetite for wood, the shipworms attack vegetable matter of every description both living and dead. The record of their ravages reveal a long list of objects from the sea, brackish waters and also from fresh waters. Roonwal (1954) reports from the 24 paraganas forest division in the Sundarbans in West Bengal, that the shipworm *Bactronophorus thoracites* attacks several species of both living and dead trees in the mangrove swamps. A recent survey (Nair & Dharmaraj 1980) in the mangrove forests connecting the Vellar and Coleroon estuaries and the adjacent coastal zones revealed the existence of several species of woodboring animals (eight species of teredinids, two of pholadids and two of sphaeromatids), including *B. thoracites* and the estuarine species *N. hedleyi*. Extensive damage has been recently reported (Dharmaraj & Nair 1980) from the aquafarms along the coasts of India .

The attack specially of shipworms is not confined to objects like wooden hulls of boats, barges and docks, but also to such objects as buoys and floats which they indiscriminately attack. Shipworms also attack floating object like cork, corky seeds and coconuts, jute or gutta-percha covers of submarine cables, ropes, plywood and lobster traps. All fixed objects of plant origin too are unsparingly attacked such as pillars of piers, wharves, stakes, poles etc., wooden water tanks in ships as well as oyster culture equipment causing huge financial loss.

Since the beginning of man's maritime activity, he has been trying every known mechanical, chemical, electrical and biological means to deter, discourage or destroy these pests. The fishermen of Bengal practice a method of suspending the boat infested by shipworms across two poles and lighting a fire beneath to destroy them. The charring of the bottom hinders further attack for a period (Wright 1864). This method is based on the fact that desiccation is fatal for these soft bodied creatures and is practised with suitable modifications in various countries. Impregnation with hydrocarbons and charring them, desiccating the infested timber, exposure to freshwater, introduction of poisons into the ambient water, fixing hard particles or completely sheathing the exposed surfaces,

exploding dynamite, electrolytic protection, electrocution, chlorinating the water etc., are some of the methods used against borers other than chemical preservatives on timber. In aquafarms chemical preservation should be tried with caution. Purushotham and Rao (1971) tentatively classified ascu, creosote-fuel oil, pure creosote and copper resinate as moderate preservatives and pentachlorophenol as poor preservative in their trials of the various preservatives.

Perhaps the most economical and effective method for checking the ravages of shipworms may be the least tried, namely biological control using the predators and parasites. Despite man's ceaseless fight shipworms with all available resources and techniques, their relentless destruction continues and a thorough reorientation in our technique of warfare has become imperative. The discovery of an effective panacea depends on a better understanding of the ecology of these specialised bivalves.

## References

1. Becker G 1958 Report to the Govt. of India on the protection of wood against marine wood-borers; *FAO Rep* 795 1-112
2. Becker G and Kohlmeyer J 1958 Deterioration of wood by marine fungi in India and its special significance for fishing crafts; *J Timb. Dry Preserv Assoc India* 4 1-10
3. Devenish-Meares R 1904 Teredo attack at depth; *Minut Proc. Cape Soc Civ Engrs* 2 6
4. Dharmaraj K and Nair N B 1979 Studies on the ecology of sphaeromatids in a tropical estuary; *Aqua Biol.* 4 49-71  
— and — 1980 Wood-boring organisms in relation to aquaculture along the coasts of India; Symp coastal aquaculture; *Mar, Biol Assoc India, Cochin*, Jan 1980  
— and — 1980a Incidence of timber boring animals in the Vellar-coleroon estuarine system; *Curr Sci* 49 486-487
5. Karande A A, Balsubramanyan K and Prema S 1968 Development of a laboratory method for bioassay of candidate toxins against teredid wood-borers; *Proc. Sym Mollusca, Mar Biol. Assoc. India*, 736-745
6. Kinne O 1963 Effects of temperature and salinity on marine and brackish water animals; I. Temperature, *Oceanogr. Mar. Biol. Ann. Rev.* 1 301-340
7. Kohlmeyer J 1963 The importance of fungi in the sea; in *Symposium on marine Microbiology* pp. 300-314 Ed C H Oppeneimer, 300-314

8. Lane C E 1955 Recent biological studies on *Teredo*, marine woodboring mollusc; *Sci Mon. N.Y* 80 286-292  
 — 1959 Some aspects of the general biology of *Teredo*, in *Marine Boring and Fouling Organisms* pp 137-144 ed D L Ray (seattle: University of Washington Press)
9. Nagabhushanam R 1959 Observations on the biology of woodboring mollusc *Bankia (Liliobankia) campanellata* (Mollusc); *Proc 1st All Indian Cong. of Zoology, Jabalpur* 2 360-363  
 — 1959a Gregarious attraction as a factor influencing the settlement of woodboring molluscs; *Proc 1st All India Cong. Zoology, Jabalpur* 2 354-355  
 — 1960 Notes on the shipworms form the Vishkapatnam Harbour ; *J Bombay nat Hist Soc* 57 362-370  
 — 1960a A note on the inhibition of marine woodboring molluscs by heavy fouling accumulation, *Sci. Cult.* 26 127-129  
 — 1961 Effect of water currents upon the rate of attack of marine borers; *J Scient ind Res* C20 66-67  
 — 1962 Seasonal settlement of molluscan wood-borers in Viskhapatnam Harbour; *Bull natn Inst Sci India* 19 131-139
10. Nair N B 1954 Shipworms of India 1. Report on ten species of shipworm from the Madras coast; *Rec Indian Mus* 52 387-414  
 — 1965 Shipworms of India II . Seven more shipworms form South India; *Rec Indian Mus* 53 261-278  
 — 1956 The development of the wood-boring pelecypod *Bankia indica*, *J Madras Univ.* 26 303-318  
 — 1957 The shipworms of India with a note on the breeding season of *Bankia indica* Nair; *J Bombay nat.Hist. Soc.* 54 344-357  
 — 1959 The marine timber boring molluscs and crustaceans of W Norway, *Publ Biol Stu Espengend, Univ Bergen Arb. Natun rekke*, 1 3-23  
 — 1960 Rate of growth of *Bankia indica* Nair, A shipworm from the Madras Coast; *J mar. Biol Assoc. India*, 2 216-220  
 — 1962 Ecology of marine fouling and wood-broing organisms of Western Normay; *Sarsia* 81-88  
 — 1963 Shipworms from the Pulicat Lake, East Coast, Madras State; *Curr Sci* 32 215-216  
 — 1965 Seasonal settlement of marine woodboring-animals at Cochin Harbour; South West Coast of India; *Int. Revue. Ges. Hydrobiol.* 50 411 422

- 1965a Seasonal settlement of fouling and boring Crustacea at Cochin Harbour; *Proc Symp. Crustacea; Mar.Biol.Assoc. India*, Cochin Part IV, 1254-1268
- 1966 Vertical distribution of marine woodboring animals in Cochin Harbour, South West Coast of India; *Hydrobiologia* 27 248-259
- 1968 The problem of timber destroying organisms along the Indian coasts, *Proc. 2nd int Cong. Mar Corrosion and Fouling Athens* (1968) 365-371
- 1979 Incidence of woodboring molluscs in the oyster farms at Tuticorin, *Mahasagar, Bull. natn. Inst Oceanogr* 12 109-114
- and — 1980 Woodboring molluscs of the Palk Bary and the Gulf of Mannar; *Mahasagar. Bull natn. Inst. Oceanogr.* 13 249-260
- and — 1983 Marine woodboring molluscs of the Lakshadweep Archipelago; *Indian J. mar. Sci* 12 96-99
- and Saraswathy M 1971. The biology of woodboring teredinid molluscs, *Adv mar Biol* 9 335-509
- 11 Purushotham A and Rao K S 1971 The first progress report of the committee for the protection of timber against marine organisms attack in the Indian coastal waters for the period 1953-1970, *J. Timb Dev Assoc India* 17 1-139
  12. Rajagopalaengar A S 1961 A new species of marine borer, *Bankia* (*Neobankia*) *roonwali* (Mollusca, Teredinidae) from India; *Sci Cult* 27 550
  13. Roowal M L 1954 *Bactronophorus thoracites* (Mollusca Teredinidae) as a pest of living trees in the Sunderbans, Bengal; *Curr. Sci.* 23 301
  - 14 Santhakumari V 1970 The life cycle of *Eutima commensalis* sp. nov. (Eutimidae, Hydromedusae); *Mar Biol* 5 113-118
  15. — and Nair NB 1969 A commensalic hydroid from woodboring molluscs; *J Nat. Hist* 3 19-33
  16. Saraswathy M and — 1969 The period of settlement of shipworms at Cochin harbour with observations on the vertical distribution of *Nausitora hedleyi*; *J Ker Acad. Biol.* 1 3-19
  - and — 1974 Growth rate of tropical estuarine shipworm *Nausitora hedleyi* Schepman (Mollusca; Teredinidae); *J. mar. Biol. Assoc. India*. Dr N K Panikker Commemoration Vol. 366-385
  - and — 1974a The influence of salinity on a tropical estuarine shipworm *Nausitora hedleyi* (Bivalvia Teredinidae); *Hydrobiologia* 44 397-411
  17. Turner R D 1966 A survey and illustrated catalogue of the Teredinidae (Cambridge, Massachusetts Museum of Comparative Zoology, Harvard University
  18. Wright E P 1864 On a new genus of Teredinidae; *Trans. Linn. Soc. Lond.* 24 451-454



**Prakash Narain Tandon** (b. 13 August 1928) did M.B.B.S. (1950), M.S. (1952) from University of Lucknow; and is F.R.C.S. (1956) from England. He is S.S Bhatnagar Fellow (1990- ), Professor of Neurosurgery, All India Institute of Medical Sciences, New Delhi (1965- ).

Tandon's researches deal primarily with problems of immediate and direct relevance to the health needs of the country. His initial researches were concerned with the delineation of the pattern of neurological diseases in India. He described new syndromes, for the first time, as also the variations in the manifestations of the well-known disorders. These led to a series of monographs on tuberculosis of the central nervous system, epilepsy in India and subarachnoid haemorrhage (standard references in these fields), head injury, its epidemiology, etiology, problems of diagnosis, delineation of special syndromes and their treatment and prognosis. His group established the significance of temporal lobe lesions, pathogenesis of growing fracture skull, and clinico-pathological profiles of brain stem haemorrhages on the basis of the largest series of such cases reported from anywhere in the world. He also carried out neuro-nuclear, neuropathological, neuroendocrinological and immunohistochemical investigations to study a variety of brain tumours. In more recent years he has focused on more basic subjects like experimental neural transplantations in rodents and sub-human primates, which have for the first time revealed the phenomenon of premature ageing and progressive degeneration of transplanted neurons.

Tandon is Fellow of Third World Academy of Sciences, National Academy of Medical Sciences, India (Vice-President, 1979), National Academy of Sciences, India (Vice-President 1989-90), Indian Academy of Sciences and Royal Society of Medicine (London), Member, Academia Eurasiana Neurochirurgica, Norwegian Academy of Science, Academy of Neuroscience, Neurological Society of India, World Federation of Neurological Surgeons (executive Committee Member, 1969-77); Corresponding Member, Scandinavian Neurosurgical Society, Honorary Member, Indo-Europe Foundation, Society of Neurological Surgeons (USA), President, INSA (1991-92). He is the recipient of Dhanwantri Award (INSA) (1986); Outstanding Alumnus Award (K.G. Medical College, Lucknow) (1987), O.P. Bhasin Award (ICMR) (1988); Honorary Surgeon to the President of India, Padma Shri (1973); Padma Bhushan (1989).

*Prakash Narain Tandon was elected to the fellowship of the Academy in 1979.*

**Address:** Professor, All India Institute of Medical Sciences, New Delhi 110029;  
**Tel:** (011) 6863112; **Res:** H-41, CSIR Complex, Maharani Bagh, New Delhi 110065; **Tel:** (011) 6834526.

## **FROM MASSES TO MOLECULES : STUDIES IN BRAIN TUMOUR BIOLOGY**

P N TANDON

*"The greatest thing in life is not the achievement but the desire to  
achieve... it is the striving that is worthwhile"*

C.V. Raman

I am deeply conscious of the honour done to me by the Academy by awarding the Shree Dhanwantri Prize. I take this opportunity to convey my grateful thanks to the Council for the same. Surgeons in the Western medical tradition have been considered as descendents of the barbers, preoccupied by the practice of the art of surgery rather than in the intellectual pursuits which characterised the physicians. In this respect, it is interesting to trace the roots of surgery in India. According to the old Dhanwantri School of Surgery, represented by the celebrated surgeon Sushruta, rightly considered the father of Indian Surgery, Lord Indira favoured Dhanwantri with the entire knowledge of Ayurveda. On being approached by a group of sages, who were moved by the human suffering, Dhanwantri agreed to admit them to his hermitage and delivered to them the science of healing. The sages selected Sushruta to be their spokesman, who it is believed recorded the very words of Dhanwantri himself (Keswani 1974). thus, it will be seen that from its very inception, surgery in our country was part of the science of healing and not just an art. Furthermore, it was founded by a sage. What a contrast to the origin of surgery in either the Unani or the so-called modern medicine! Critically evaluating the characteristics of the founders of Indian medicine, Keswani (1974) concludes that while Charaka, in his writings, had a combined role of a moralist, philosopher and above all a physician, Sushruta tried to cast off whatever shackles of priestly domination remained at his time, and created an atmosphere of independent thinking and investigation. It is in this tradition that I have chosen for presentation today those aspects of our work, which have been guided by the spirit of scientific enquiry rather than the ones concerned with the practice of the art of neurosurgery.

Till recently, the neurosurgeons were mostly preoccupied with the diagnosis and surgical excision of masses (tumours, granulomas, abscesses, clots) impinging upon or involving the brain and the spinal

cord. In the earlier years, the diagnosis was based on clinical assessment alone. It often led to opening the skull at the wrong place with disastrous consequences. The quest for a more reliable test to correctly localise the masses resulted in successive introduction of ventriculography, pneumoencephalography and cerebral angiography. These were all invasive procedures, invariably producing some discomfort to the patient and occasionally a serious complication. Nevertheless for nearly 50 year, these remained the sheet-anchor for localising the intracranial space-occupying lesions, while myelography was introduced for diagnosing the intraspinal masses. For these techniques to establish the diagnosis, the lesion should have already acquired a sizeable mass. The diagnosis was based on indirect evidence of space occupation, the nature of the mass generally remained uncertain. Electroencephalography (EEG) initiated by Berger (1929) provided the first non-invasive method, but it could at best be used as a screening procedure, precise localisation being seldom possible. Isotope encephalometry, also non-invasive, was far more precise, specially for tumours in certain locations and histological types. Nevertheless, it had its own limitations. It was only with the invention of the CT scan by Hounsfield in 1973, that a completely non-invasive, totally safe and remarkably precise diagnostic tool became available, much to the relief of the patients and neurosurgeons alike. Besides revealing much smaller lesions, even a few millimeters in size, a direct demonstration of the tumour, often indicated its nature. Thus, it was possible for us to establish the image morphology of intracranial tuberculomas, diagnose these when only few millimeters in size (the so-called Microtuberculomas) and evolve a safe therapeutic regime (Bhargava & Tandon 1980 a, b, Tandon 1983, Tandon & Bhargava 1985). All these advances, improving the safety and reliability of diagnosis of a mass lesion only provided information of a physical nature. This was undoubtedly valuable from a purely technical standpoint but contributed little to the basic understanding of the pathophysiology of the lesion, without which the real challenges of brain tumours can not be met. Already in 1970, we had realised that to meet these challenges "a collaborative study involving the neurosurgeons, pathologists, molecular biologists and immunologists, is the only hope to find out counter moves to fight "the utterly relentless antagonist"—as the brain tumours were characterized by Cushing (Tandon 1970).

Two recent technological advances—(Nuclear) Magnetic Resonance Imaging (MRI) supplemented with *in-vivo* spectroscopy and the Positron Emitting Tomography (PET)—promise to provide additional information of

far greater functional importance in addition to the physical image of the lesion. The surgeon for the first time starts to move not only from the large masses to small lesions but to the molecules constituting it. Unfortunately, these facilities are still not available in India for us to make any personal contributions. Nonetheless, these open such new vistas for the study of the disease processes themselves, rather than just their image, that I may be permitted to spend a little time before describing our own march from the masses to the molecules.

Magnetic Resonance (MR) is a phenomenon that was discovered independently by Purcell et al. and Block in 1946, for which they received the Nobel Prize in Physics in 1952. Eversince it has been used for chemical analysis, but its translation from spectroscopy to imaging, from chemical compound to intact biological systems and from there to patients affected by disease process is less than a decade old. MRI is largely based on the response of certain atomic nuclei in a magnetic field to electromagnetic energy at a radio frequency. It provides cross-sectional images of the body which consist essentially of distribution density maps of the mobile protons in cellular water and lipids. In certain types of lesions and in some locations the images provided by MR are much superior to the ones obtained from X-ray CT. However, it is not the imaging capabilities of MR, but its potentialities for *in-vivo* spectroscopy permitting the study of chemical state of the tissue constituents as also bioenergetics, with its promise for providing new perspectives on the etiology and pathogenesis of disease process, which distinguishes NMR from most other diagnostic radiology techniques. Already  $P^{31}$  *in-vivo* spectroscopy has provided valuable information on the metabolic muscle diseases, tumour response after chemotherapy and for stroke patients.

Positron Emitting Tomography (PET Scan) is yet another development which has made it possible to study the chemical physiology of the brain *in-vivo*. Still comparatively much inferior to either CAT scan or MRI as far as the image morphology is concerned, it is by far the most sensitive investigation to study the regional cerebral blood flow and blood volume, regional cerebral metabolic rate for oxygen and glucose, regional oxygen extraction fraction, and pharmacokinetics of anti-tumour and other drugs acting on the nervous system. Thus, it has made it possible to measure biochemical changes during the evolution of cerebral ischaemia, focal epilepsy, brain tumours, etc., in intact brain. An apt comparison between CAT and PET scans has been provided by Dr William Feindel, one of the pioneers in this field, who stated, "the CAT gives you a picture



of the building, the PET gives you a picture of what's going on inside the building—the people, the plumbing, the electricity." In true sense, PET provides an image of the events taking place at a molecular level in a given mass of tissue in the living brain.

In absence of such facilities, the urge to enhance our understanding of the masses which we commonly encountered in our day to day care of the victims of brain tumours turned us to the limited resources of our own laboratories. Utilising the new techniques of modern biology, like radio-immuno-assay (RIA), use of tumour markers and tissue culture, it has been possible to study some of the functional characteristics of brain tumours at the cellular and molecular level. The following account briefly documents our efforts in this direction.

### PITUITARY TUMOURS

Nearly 10 years ago, while analysing our experience in the management of a large series of brain tumours during the previous decade, the satisfaction derived from our technical successes was tempered by a feeling of inadequacy born out of prevailing limitations of knowledge of the more basic characteristics of biological and functional significance (Tandon 1977). To take one example, that is of pituitary tumours, we had already achieved an operative mortality of less than 5%, notwithstanding the very advanced stage of the disease, we are called upon to treat. This was comparable to the best reported from anywhere in the world. It may be pertinent to quote from this paper, published in 1977, "The Neurosurgeons are happy to relieve the symptoms of pressure on the optic pathways, ensure improvement in vision, but are we able to ensure a normal human being?" It was obvious that the removal of the mass only provides relief from the mechanical component of the lesion, that is pressure on the neighbouring structures. This limited approach, however, satisfying to the surgeon's ego, falls far short of the overall needs of the patient. This prompted us to initiate a multi-disciplinary study to improve our understanding of the basic biological and patho-physiological factors determining the morbidity of these patients. Thus, besides the routine clinical and radiological evaluation, a series of investigations utilizing the latest sophisticated techniques have been carried out for this purpose. Professor Subimal Roy and his colleagues (Drs Chitra Sarkar and Meera Mathur from the Department of Pathology) and Professor Kochupillai and his team (specially Dr Nandita Gupta from the Department of

Endocrinology) have been responsible for the pathological and neuroendocrinal aspects of the study.

A critical review of our cases of pituitary adenomas confirmed the growing feeling that the time-honoured classification of pituitary adenomas into chromophobes, acidophil and basophil did not reflect the hormonal status of the patient. Thus, patients with histologically identical tumours, e.g. a chromophobe adenoma, could manifest with pituitary hypofunction or hyperfunction including acromegaly and/or hyperprolactinaemia (Prakash *et al.* 1974 Mohanty *et al.* 1977). Electronmicroscopy of a series of such tumours revealed that a much better assessment of the secretory capacity of the tumour could be made by this investigation (Ray 1977, 1983). An evaluation of the cytoplasmic organelles like rough endoplasmic reticulum, golgi apparatus, mitochondria and secretory granules could differentiate a secretorily active tumour from an inactive one and also give an idea of the storage capacity of the tumour cells. This helped us to have a better understanding of the functional status of some of these patients. However, contrary to the earlier claims, it was obvious that even EM fails to consistently reveal the nature of the hormone secreted or stored, which in turn determines the clinical status of the patients. Unlike the normal pituitary gland, size of the granules does not always indicate nature of the hormone in the adenoma. To have a better understanding of the biological behaviour of these tumours, it was considered necessary to have a correlative study including the clinical picture, light and electronmicroscopic findings with serum hormone levels and the hormone content in the tumour tissue itself. RIA and immunohistochemical investigations were utilised for this purpose. Some of the tumours were cultured and their hormone secretion studied *in vitro*. It may be pointed out that all the tumours studied in this series were large or so called macroadenomas. Without going into the details of these studies, which are documented elsewhere (Roy *et al.* 1984, Chowdhry *et al.* 1986, Gupta 1984, 1985, 1986), it is worthwhile summarising some of the significant conclusions:

- (i) The time-honoured classification of pituitary adenomas into chromophobes, acidophils and basophils based on the routine light microscopy (H & E or PAS orange O. stain) is inadequate to consistently reflect the functional characteristics of the tumour.
- (ii) Electron microscopy clearly distinguishes tumours which are secretorily active from the inactive, and it may provide an indication

as to the possible nature of the hormone produced in some but not in most cases.

- (iii) Immunohistochemistry clearly demonstrates the nature of the hormone present in the tumour. However, it does not automatically imply that the hormone so visualised is being secreted and is contributing to the clinical picture. This may be due to: (a) small number of cells secreting the hormone, (b) lack of normal transport mechanism, (c) block at the receptor level in the target organ, or (d) the hormone being biologically inactive.
- (iv) RIA estimation of a hormone in the tumour itself appears to be more sensitive in detecting the presence of the hormone but suffers from the same limitation as immunohistochemistry *vis-a-vis* its clinical significance.
- (v) Serum hormone levels provide a much better correlation with the clinical picture. Nevertheless, there are exceptions even to this, for example patients with hyperprolactinemia may or may not have galactorrhoea.

It is, therefore, obvious that each of these investigations provides valuable, complementary information, all of which is essential to understand the biological behaviour of these tumours. It is only with such comprehensive knowledge at the cellular and molecular level that the neurosurgeon's capabilities, not just to relieve pressure on the adjacent structures but to treat the man as a whole, will depend.

The most unique observation of this study was the detection of immunoreactive prolactin (PRL) in concentration statistically higher than in normal human pituitaries in all the tumours examined. There was no statistically significant difference in the PRL concentration in the tumour tissue from four groups of patients studied, i.e. acromegalics with or without hyperprolactinaemia as compared to the concentration in non-acromegalics with or without hyperprolactinaemia. It is, therefore, logical to conclude that at least as far as the macroadenomas of the pituitary are concerned PRL appears to be a tumour marker. Immunoreactive growth hormone (GH) was also detected in all these tumours. This was, however, lower than in the normal pituitary gland ( $p < 0.001$ ). There was a significantly higher concentration of GH in the tumour tissue obtained from the acromegalics compared to the non-acromegalics ( $p < 0.001$ ).

The significant difference in the PRL : GH ratios observed between the four groups of patients mentioned above is reflected in the spectrum of their clinical and endocrine features (Gupta et al. 1986). Thus, the majority of females belonging to the age group up to the 3rd decade, tend to harbour adenomas that preferentially produce (as reflected in the tissue PRL : GH ratio) and secrete GH to attain high circulating levels. The female-dominated first peak of prevalence of pituitary adenomas would thus seem to be due to the preponderance of this type of adenomas. However, the pure prolactinomas (group 3), which produce and secrete PRL, seem to occupy an intermediate position from the point of view of age-related prevalence.

In contrast, endocrine inactive adenomas were a set of tumours which predominantly tend to occur in patients above 30 years, and constituted 20% of all the macroadenomas. They accounted for the second peak of age-related prevalence dominated by males. The demonstration of significant concentration of PRL immunoactivity in the tissue of such adenomas, suggests that these tumours though producing PRL, were poor secretors of the hormone, thus showing dissociation between production and secretion of PRL. In this context, it is interesting that 1/3 of PRL immuno-reactivity measured in such tumours was "big-big" PRL, indicating *pari-passu* abnormalities of cytoplasmic processing of PRL precursor peptide, along with abnormalities of secretion of the hormone.

The demonstration of co-existence of PRL and GH immuno-reactivity in all the macroadenomas, as well as their interesting ratio-relationship in different groups suggesting it to be a function of the degree of dedifferentiation of a tumour, may well have very interesting implications from a molecular biological point of view. Thus, both PRL and GH are recognised to be related molecules which have evolved from a common precursor molecule by gene duplication and random mutation through evolution. Being related molecules, PRL and GH may well have related regulatory mechanisms for their expression. In this context, it is interesting to recall that normally GH is under stimulatory control and PRL expression is under inhibitory control of the hypothalamus. However, there are a number of reports in the literature to show that these reciprocal regulatory mechanisms are thrown out of gear in patients with pituitary neoplasms (Leavens et al. 1977, Refetoff et al. 1979, Reichlin 1980). While the precise mechanisms of these abnormal responses remain unknown, they clearly indicate dislocation in the normal cytoregulatory process of GH and PRL secretion, in a neoplastically transformed cell.

These dislocations have been interpreted as indicative of cellular dedifferentiation by some authors. De-differentiation of a normal somatotroph is believed to result in its acquiring lactotroph like characteristics. These considerations make it likely that PRL production may well be the attribute of a de-differentiated somatotroph (Martin 1979, Halmi 1982).

The results of our study seem to suggest that expression of PRL immunoreactivity is a functional concomitant of neoplastic proliferation and de-differentiation. There are several reports in the literature linking PRL and cell proliferation in the pituitary and non-pituitary tissues. These considerations impel one to recall the relationship between oncogenes and neoplasia on the one hand, and growth factors which are products of oncogenes on the other. In the light of these, it remains a tantalising possibility that PRL is indeed the product of an oncogene.

The existence of more than one hormone in the tumour, as revealed by the RIA technique has further been confirmed by our more recent immunohistochemical investigations (Roy et al. 1984, Chowdhury et al. 1986). While both PRL and GH were detected by RIA in varying proportions in all these tumours, immunohistochemistry revealed it in many but not all of these. Two or more hormones were demonstrated in 51% of 95 tumours studied immunohistochemically. This difference may be due to the very high sensitivity of RIA which can detect picogram quantities of PRL and GH in the tissue. Tissue culture of tumours from 12 patients revealed both PRL and GH detectable by RIA in the supernatant fluid in all of them. These studies clearly indicate that many, if not all macroadenomas of the pituitary, arise from pluripotent stem cells and can show differentiation into one or more cell lines.

While the occurrence of pleurhormonal pituitary adenoma has been recognised for some time (Zimmerman et al. 1974, Lamberts et al. 1979, Horvath & Kovacs 1980, Ryder et al. 1980, Kovacs 1986), their true incidence is not known. All the reported series are based on immunohistochemistry. Ours is the first study of hormone estimation in the tumour using RIA.

The interesting question of the source of multiple hormones in a given tumour needs further elucidation. There is evidence to suggest that such polyhormonal tumours may be monomorphic, composed of one cell type which produces more than one hormone, or plurimorphous,

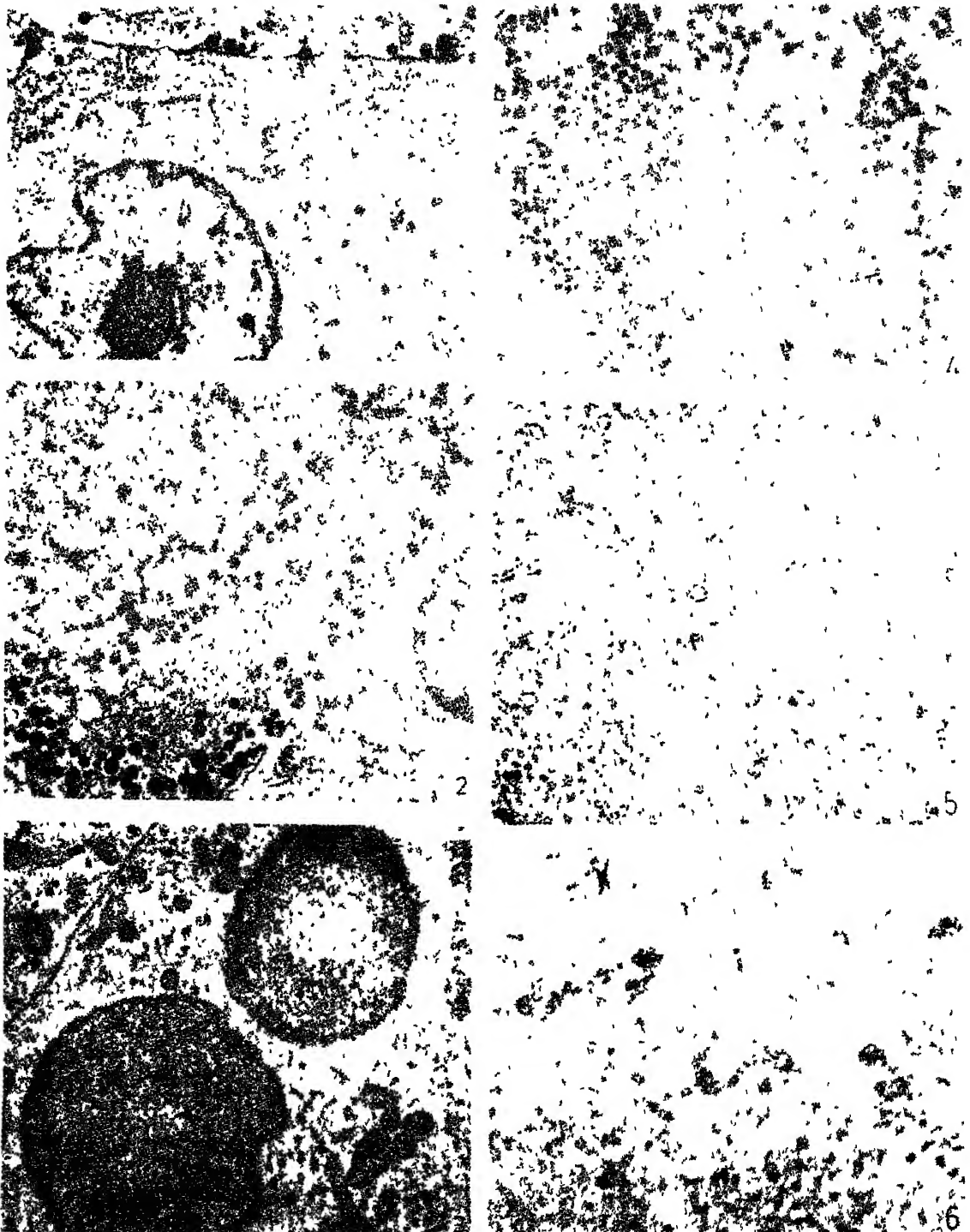
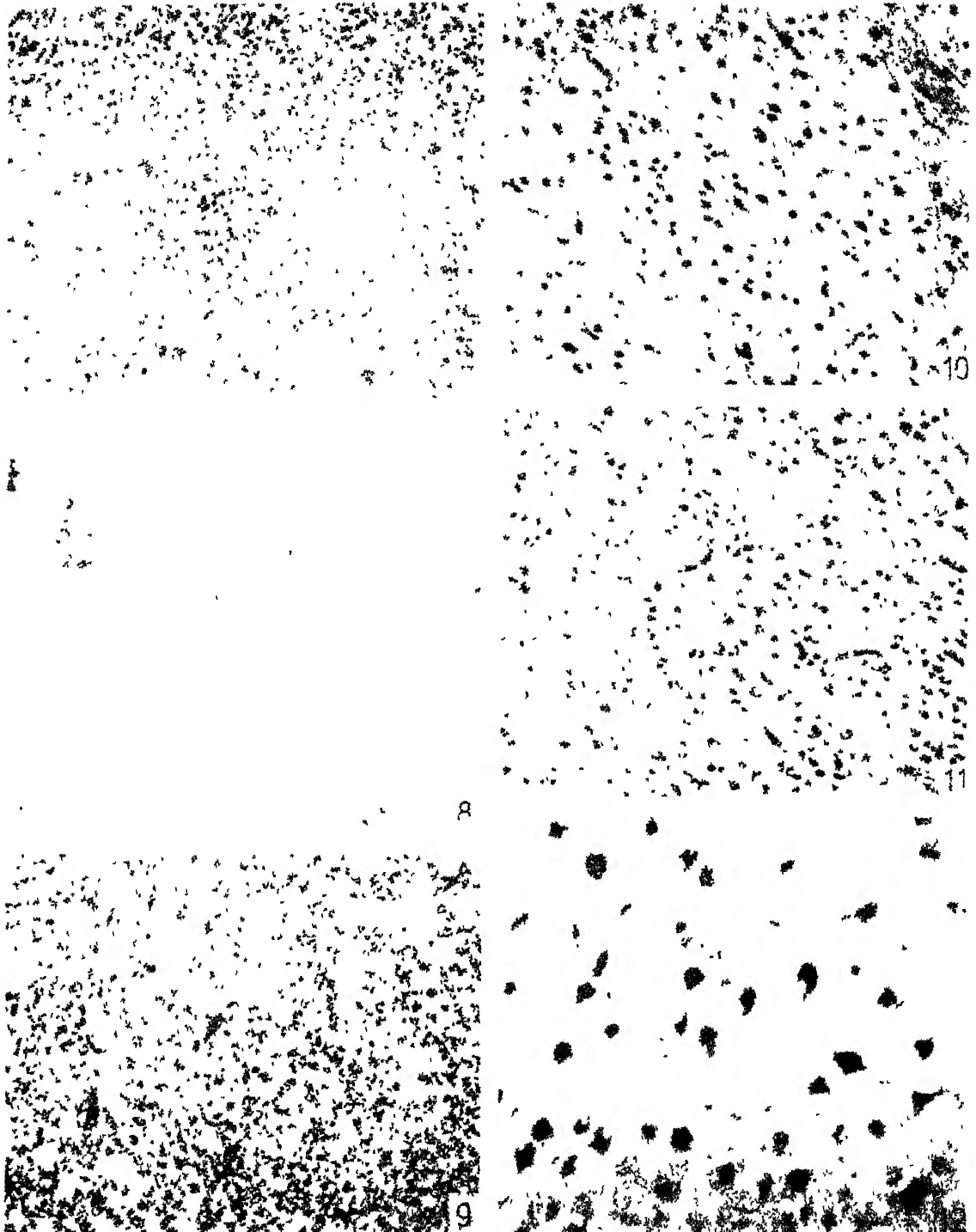


Fig 1-6 1. Electronmicrograph showing secretorily inactive cells of a pituitary adenoma. Note the paucity of RER and golgi apparatus (X32,400); 2. Electronmicrograph showing secretorily active cells of a pituitary adenoma. Note prominent RER and golgi apparatus and a few secretory granules (X399,460); 3. Electronmicrograph showing a granulated binucleate cell of pituitary adenoma with prominent RER (X53,040); 4. Photomicrograph showing prolactin-positive cells (PAPX300); 5. Photomicrograph showing growth hormone positive cells in the same tumor (PAPX300); 6. Photomicrograph showing ACTH positive cells in the same tumor (PAPX300).



**Fig 7-12** 7. Photomicrograph showing undifferentiated cells of a primitive neuroectodermal tumor. Note sheets of small dark staining cells (H&EX300); 8. Photomicrograph showing GFAP positive cells in a primitive neuroectodermal tumor. Note the astrocytic processes. (PAPX500); 9. Photomicrograph showing neurofilament positive cells in primitive neuroectodermal tumor indicating neuronal differentiation (PAPX300); 10. Photomicrograph showing astrocytic area in a mixed glioma. Note the fibrillary background (H&EX3000); 11. Photomicrograph showing oligodendroglial area in a mixed glioma (H&EX300); 12. GFAP-positive oligodendroglial cells in a mixed glioma (PAPX500).

consisting of two or more distinct cell lines each producing one hormone (Corenblum et al.1976, Horvath & Kovacs 1980, Halmi 1982).

Reviewing the existing literature and his own extensive work, Kovacs (1986) concluded, "The cytogenesis of plurihormonal pituitary adenomas is obscure." We are currently engaged in studying this feature with the help of immuno-electronmicroscopy and tissue culture.

The new knowledge added by these investigations is not purely of academic interest, but even at this stage it has practical implications for the management of patients with pituitary adenomas. It is now established that bromocriptin, not only counteracts hyperprolactinaemia but also regresses even large prolactinomas. Likewise, somatostatin analogues have been demonstrated to be beneficial for patients with growth hormone secreting tumours. However, the role of these drugs in the management of so-called non-functional pituitary adenomas, i.e. those without hyperprolactinaemia or hypersomatotropism, but with increased levels of these hormones in the tumour as demonstrated by RIA or immunohistochemistry, is still not known. On the basis of the current observations it would be worthwhile instituting a therapeutic trial on these lines. This may be specially useful in patients with residual or recurrent tumour for which the tumour hormone contents have already been established at initial surgery. Another important question raised by these observations is the role of multitherapy for pleurihormonal tumours. It is intended to pursue these leads in selected patients.

### NEUROECTODERMAL TUMOURS

Another area of abiding interest in neuro-oncology, since the classification of brain tumours by Bailey and Cushing (1926), has been their histogenesis. The original classification presumed that the primitive neuroepithelial cell differentiates into several distinct cell lines; astrocytic, oligodendroglial, ependymal; the type of tumour that develops depends upon the cell line affected and its stage of differentiation. A variety of theories have since been postulated. Dealing with such large numbers of brain tumours as we do, dissatisfied by the existing classification which not only failed to account for all the tumours we saw but had obvious limitations with regards to predicting their biological behaviour, we have over the years continued to study our own material in search of finding some explanation for the yet ill-understood observations. Towards this end, we have utilised the new biological tools as these became available



during the last years. For the purpose of this presentation our studies on two such groups of tumours which provide new insight into oncogenesis of the glial tumours have been selected.

### PRIMITIVE NEUROECTODERMAL TUMOURS

A group of brain tumours characterised by poorly differentiated neuroepithelial cells, whose histogenesis and classification have been the subject of much controversy, were investigated to resolve these controversies. These include tumours variously described as medulloblastoma, cerebral or central neuroblastoma, medulloepithelioma, pineoblastoma, olfactory neuroblastoma, ependymoblastoma and primitive neuroectodermal tumours. On light microscopy many of them are indistinguishable from each other. Often the label is given based on the knowledge of their location rather than tissue morphology (Rorke 1983, Rorke et al. 1985, Rubinstein 1985). The commonest tumour in this group is medulloblastoma. Prompted by an observation at surgery, some years ago, it was decided to study the differentiating potentials of these tumours. Ultrastructural studies by Roy et al. (1977) confirmed our suspicion that this tumour has capabilities to differentiate both towards glial, predominantly astrocytic, and neuronal lines. This was further confirmed by immunohistochemical studies utilising antibodies for glial fibrillary acidic protein (GFAP) (Roy et al. 1984, Chowdhry et al. 1986). Around the same time we were impressed by evidence of differentiation and de-differentiation in some of the recurrent tumours when comparing the histology of the tumour removed at the initial surgery with that for recurrence (Roy et al. 1980). With the availability of a variety of specific markers for various neuroectodermal cells it was decided to study the whole group of these tumours for evidence of their histogenesis. Besides electronmicroscopy, immunohistochemical studies were carried out utilising antibodies against GFAP, neurofilaments and S100 protein (Sarkar et al. 1986).

The collective evidence from all these studies reveals that most of these tumours have origin from a pleuripotential cell with possibilities to differentiate into glial and neuronal lines. It would thus be desirable to collectively call them primitive neuroectodermal tumours and direct our attention to discover the pathogenetic mechanism responsible for the aberration for the group as a whole rather than each individual tumour. The molecular events which determine the direction and degree of change in a given tumour are not known at the present stage. However, it is

reasonable to conclude that neuroectodermal tumours form a spectrum exhibiting varying degrees of differentiation and de-differentiation, which may have origin in a totipotent cell. There may thus be no need to postulate a specific cell of origin for any given neoplasm in the brain. This finds further confirmation in the next group of tumours, i.e. mixed gliomas, studied by us. The clinical significance of these observations is not clear at this stage. Further studies in this regard, specially a correlation between their morphology, biological behaviour and response to therapy would be rewarding.

### ***Mixed Gliomas***

During a critical analysis of a large number of patients with supratentorial gliomas treated in our department during the last two decades, we were impressed by a group which appeared to constitute a distinct entity, clinically and pathologically. However, this group seems to have been generally overlooked by most neurosurgeons and neuropathologists (Hart et al. 1974). This was pointed out during the National Seminar on Neuro-oncology held at Bangalore in 1979 (Tandon 1979). Since that time we have carefully studied nearly one hundred such cases. On the basis of our investigations, there is enough evidence to suggest that this group of tumours merits a distinct place in the nosology of brain tumours. We have preferred to use the term mixed glioma for this group of tumours, because at least two types of glial cells, usually astrocytes and oligodendrocytes, are seen to constitute distinct areas of the tumour. However, others have called it oligoastrocytoma (Herpers & Budaka 1984).

These tumours consisted chiefly of areas of astrocytoma and oligodendroglioma. In some cases ependymal areas were also seen. Endothelial proliferation was present in some of the cases and foci of calcification were seen in 18 out of 25 cases. The astrocytic areas had fibrillary background and both cells and fibres were intensely positive for GFAP. The ependymal area was GFAP positive. In the oligodendroglial areas of these tumours, two types of cells could be seen. One type was the classical box like cells with a clear halo around the nucleus. These cells showed variable degree of GFAP positivity. Some were negative, others showed a blob in the perinuclear region while still others showed a thin rim of positivity at the periphery. The second type of cells were small with eosinophilic cytoplasm and eccentric nuclei. These were intensely GFAP positive and resembled small gemistocytes. These gemistocytes were intimately admixed with the vacuolated (clear) cells.

A similar pattern of GFAP reactivity as seen in the oligodendroglial areas of the mixed glioma was observed in all the fifteen cases of pure oligodendrogliomas studied.

Beyond just the nosological interest, these tumours provided an opportunity to study some biological characteristics which are of interest not only in respect of oncogenesis but also ontogenesis of the astrocytes and oligodendrocytes. GFAP has been recognised as a characteristic protein for astrocytes. Stray reports in the literature had suggested the presence of this protein in typical oligodendroglioma cells in grade III tumours (Vander Meulen et al. 1978), in mouse glioma cell lines capable of producing myelin-related glycosphingolipids specific to oligodendrocytes (Neskovic et al. 1981), and myelin forming glia during normal development (Choi & Kim 1984). However, using GFAP antisera, conflicting results have been described by various investigators with regard to the existence of GFAP immunoreactive neoplastic cells other than astrocytes in such tumours. GFAP-positive cells have been reported in oligodendrogliomas by De Armond et al. (1980), Ishida et al. (1982) and Meneses et al. (1982) while Eng and Rubinstein (1978), Tascos et al. (1982), Velasco et al. (1980) could not confirm it. Herpers and Budka (1984) in a study of 50 oligodendroglioma and 16 oligoastrocytomas found GFAP-positive cells in 50% of the former but in only two of the latter. Transitional form was seen in 32% of pure oligodendroglioma and in only one of the mixed oligoastrocytoma.

The similar pattern of GFAP reactivity in the oligodendroglial areas of the mixed glioma and in the pure oligodendrogliomas indicates that these tumours possibly arise from a single precursor cell which has the potential to differentiate along both cell lines. Raff et al. (1983) have shown that fibrous astrocytes (or Type II astrocytes) and oligodendroglia develop from a common progenitor cell. Such a common progenitor cell may acquire neoplastic potential and proliferate to give rise to this group of tumours. Depending upon the degree and the predominant line of differentiation, the stage at which the patient presents to us, as also limitation posed by sampling of a biopsy, one may see a mixed glioma or a pure oligodendroglioma. Choi and Kim (1985) and Ogawa et al. (1985) have shown that immature oligodendroglial cells in developing human foetal spinal cord and *in-vitro* transiently express GFAP. So, the GFAP-positive oligodendroglia may suggest the return of a foetal behaviour by some neoplastic oligodendroglia. The gemistocyte-like GFAP-positive

cells may represent transition forms between astrocyte and oligodendroglia.

This brings us to the question as to what determines the degree of differentiation in these tumours if they do arise from a common progenitor cell. This may be related to gene expression. It has been shown that a substance 5-Azacytidine, which is known to change gene expression, when added to *in-vitro* cultures of fibroblasts causes their differentiation into chondrocytes, fat cells and muscle (Riggs 1983). Another example is experimental tumorigenesis, where giving the same carcinogenic agent in identical situations gives rise to a spectrum of glial tumours. This has been attributed to difference in gene expression by Rubinstein (1976). Of course, at this stage the factors which determine such gene expression are not known, but the real understanding of the biology of these tumours would come through following these leads.

During the course of daily professional work, the sacred duty of treating those who have trusted their lives to us, there have been moments of doubts and inadequacy about our knowledge concerning their ailment. Continuous review of the accumulating experience, its critical evaluation in the light of emerging knowledge, timely discussions with capable colleagues from allied disciplines, have resulted in delineating specific questions and planning appropriate investigations. The results of some of these investigations, the collective efforts of a number of colleagues, who were ever willing partners in this pursuit have been presented.

There is no doubt that improved methods of diagnosis and technically high standard of surgery have continuously improved our capabilities in providing relief to patients with masses (tumours) in the brain, basic understanding of their biology is essential to meet their challenges appropriately. This is what we have ventured. It has been not only a source of continuous intellectual stimulation for me, which provided an escape from the rigors of daily professional work, but a way to elevate it to an exciting occupation. In this process, if it has been possible to resolve some of the existing doubts, and add some new knowledge with the potentials for improving the care of our patients, it is an added reward.

Let me conclude by quoting one of my teachers, Dr Wilder Penfield, who combined the art of surgery with a spirit of scientific enquiry most admirably; "Thus I say, that for each of the arts there is a purpose beyond the skill and in that purpose there are eternal values for society."

## References

1. Bailey P and Cushing H 1926 A classification of the tumours of the glioma group on a histogenetic basis with a correlated study of prognosis (Philadelphia: J B Lippincott Co)
2. Berger H 1929 Uber das Electrnkephalogram des Menschen Arch. f; *Psychiat Nervenkr.* **87** 527-570
3. Bhargava S and Tandon P N 1980a Intracranial tuberculoma—A CT study. *Brit J Rad* **53** 935  
— and — 1980b CNS tuberculosis: Lessons learnt from CT studies' *Neurol (India)* **28** 207-212
4. Block F 1946 Nuclear induction; *Phys Rev* **70** 460-474
5. Choi B H and Kim R C 1984 Expression of glial fibrillary acidic protein in immature oligodendroglial; *Science* **223** 497-499
6. Chowdhury C, Roy S, Gupta N, Kochupillai N and Tandon P N 1984 Chromophobe adenomas of pituitary gland—A light microscopic and immunohistochemical study Presented at the *Second International Symposium on Biology of Brain Tumours*, London, *J Neurooncology (Abstr 33)*  
— , — and Tandon P N 1985 Differentiating medulloblastomas — An immunohistochemical study using glial fibrillary acidic protein (In press)
7. Corenblum B, Sirek A M T, Horvath E, Kovac K, Ezrin C 1976 Human mixed somatotrophic and lactotrophic pituitary adenomas; *J Clinic Endocrine Metab* **42** 857-863
8. De Armond S J, Eng L F and Rubinstein L J 1980 The application of glial fibrillary acidic protein immunochemistry in neurooncology; *Pathol Res Proc* **168** 374-394.
9. Eng L F and Rubinstein L J 1978 Contribution of immunohistochemistry to diagnostic problems of human cerebral tumours; *J Histochem Cytochem* **26** 513-522.
10. Gupta N, Kochupillai N, Godbole M M, Goulata R K and Tandon P N 1984 Prolactin and growth hormone immunoreactivity in pituitary macroadenomas. *Excerpta Medica International Congress Series* **652** Abstr 1030  
— 1985 Immunoreactive prolactin, growth hormone, leuterinizing hormone and thyroid-stimulating hormone in pituitary macroadenomas; *Ph.D. Thesis, AIIMS*  
—, Kochupillai N, Tandon P N and Ahuja M M S 1986 Neoplastic spectrum of giant pituitary macroadenomas. (*Under Communication*)
11. Halmi N S 1982 Occurrence of both growth hormone and prolactin immunoreactive material in cells of human somatotropic pituitary adenomas

containing mammotropic elements; *Virchow Archives* (Pathologica et Anatomica) **398** 19-31

12. Hart M C, Petito C K and Earle K M 1974 Mixed Gliomas; *Cancer* **33** 134-140
13. Herpers M J H M and Budka H 1980 Glial fibrillary acidic protein (GFAP) in glial tumours. Glufibrillary oligodendroglioma and transitional oligoastrocytoma as subtypes of oligodendroglioma; *Acta Neuropathol* (Berl) **64** 265-272
14. Horvath E and Kovacs K 1980 Pathology of the pituitary gland; in *Pituitary Diseases* eds C Ezrin, E Horvath, Kaufman K Kovacs (Weiss MH: CRC Press, Boca Raton) 1-83
15. Hounsefield GN 1973 Computerized transverse axial scanning (tomography) I Description of System; *Brit J Radiol* **46** 1016 - 1022
16. Ishida, Y, Takahasi K and Nakazato Y 1982 Immunohistochemical and electronmicroscopic studies of experimental and human oligodendrogliomas; *Proc 11th int. Congr. Neuropath Vienna* p. 181 (Abstr. 1-42)
17. Keswani N H 1974 Medical heritae of India; in *The Science of Medicine and Physiological Concepts in Ancient and Medieval India* ed N H Keswani; Secretary General XXVI International Congress of Physiological Sciences (New Delhi AIIMS)
18. Kovacs K, 1986 Pathology of Pituitary adenomas; *Neurology* (India) **34** 147-154
19. Lamberts S W J, Klijn J G M, Kwa G B and Birkenhager J C 1979 The dynamics of growth hormone and prolactin secretion in acromegalic patients with "mixed" pituitary tumours, *Acta Endocrinol* **90** 198-210
20. Leavens M E, Samaan N A, Jesse R H and Byers R M 1977 Clinical and Endocrinological evaluation of 16 acromegalic patients treated by transsphenoidal surgery, *J Neurosurg* **47** 853
21. Martin J B 1979: In *Clinical Neuroendocrinology — A pathophysiological approach* ed. G. Tolis, F Labrie, J B Martin, F Naftolin (New York: Raven Press) p. 269.
22. Meneses A C O, Kepes J J and Sternberger N H 1982 Astrocytic neoplastic oligodendrocytes; *J Neuropathol Expl. Neurol* **41** 368
23. Mohanty S, Tandon P N , Banerji A K, Prakash B, and Roy S 1977 Pituitary adenoma clinico-pathological study of 70 cases; *J Surg Sciences* **13** 32-35
24. Neskovic N M, Rebeal G, Harth S and Mandel P J 1981 Biosynthesis of galactocerebrosides and glucocerebrosides in glial cell lines; *J Neurochem* **37** 1363-1370
25. Ogawa H, Sato Y, Takeshita I and Kitamura K 1985 Transient expression of glial fibrillary acidic protein in developing oligodendrocytes *in vitro*; *Develop Brain Research* **18** 133-141

26. Prakash B, Tandon P N and Banerji A K 1974 Tumours in and around sella Study of one hundred cases; *Phronesis* 3 141-152
27. Purcell E M, Torrey H C and Pound R V 1946 Resonance absorption by nuclear magnetic moments in a solid; *Phys Rev* 69 37-38
28. Raff M C, Miller R H and Noble M 1983 Glial progenitor cell that develops into an astrocyte or an oligodendrocyte depending on culture medium; *Nature* 303 390-396
29. Refetoff S, Frank P H, Roubebusch C and DeGroot L J 1979 Evaluation of pituitary function; in *Endocrinology* Vol. I ed. L J DeGroot, G F Cahill and L Martini (New York: Grunne & Stratton) p. 175
30. Reichlin S 1980 In *Pituitary Adenomas*, ed K D Post, I M D Jackson and S Reichlin (New York: Plenum Publ. Corp.) p. 29
31. Riggs A D 1983 5-Methylcytosine gene regulation and cancer; *Adv, Canc Res* 40 1-30
32. Rorke L B, Grilles F H, Davis R L and Becker L E 1985 Revision of the World Health Organisation. Classification of Brain Tumours for Childhood Brain Tumours; *Cancer* 56 1869-1886
33. Rorke L B 1983 The cerebellar medulloblastoma and its relationship to primitive neuroectodermal tumours; *J Neuropath Exp Neurol* 42 1-5
34. Roy S 1977 Ultrastructure of chromophobe adenomas of the human pituitary gland; *J Pathol* 122 219-224  
 — 1983 Ultrastructure of Pituitary Adenomas with particular reference to chromophobe adenoma, *Progress in Neuropathology* 5 223-259  
 —, Chowdhury C, Kochupillai N and Tandon P N 1984 Chromophobe adenoma of pituitary gland — a light microscopic, electron microscopic and immunohistologic and immunohistochemical study. Presented in the *Second International Symposium on the Biology of Brain Tumours* held in London (October 24-26); *J Neuroonc Abstract* 34  
 —, — and Tandon P N 1986 Differentiating medulloblastomas — A light microscopic, electron microscopic and immunohistochemical study (under communication)  
 —, Shanker S K and Tandon P N 1977 Differentiating medulloblastomas; *Neurology (India)* 25 219-225  
 —, — , — 1980 Differentiating and de-differentiating potentials in cerebral neuroectodermal tumours; *Neurology (India)* 28 17-22
35. Rubinstein L J 1976 Current concepts in neurooncology; *Adv Neurol*, 15 1-25

- 1985 Embryonal central neuroepithelial tumours and their differentiating potential. A cytogenetic view of a complex neuro-oncological problem; *J Neurosurg.* **62** 795-805
36. Ryder D A, Horvath E, Kovacs K 1980 Fine structural features of secretion in adenomas of human pituitary gland; *Arch Pathol Lab Med* **104** 518-526
37. Sarkar C, Roy S and Tandon P N 1986 Primitive neuroectodermal tumours of the central nervous system. An electron microscopic and immunohistochemical study (Under Preparation)
38. Tandon P N 1970 Challenge of brain tumours *Bull, AIIMS* **13** 115-118
- 1977 Travails of neurosurgery in developing countries, *Proc Inst, Neurol, Madras*, **7** 79-89
- 1981 Classification of CNS tumours Controversies from surgical point of view. Neurooncology. *Proc National Seminar NIMHANS* 73-78
- 1983 Neurotuberculosis - A Continuing Challenge, *Ann natn Acad, Med Sci* **19** 84-93
- 1986 Symposium on Pituitary Tumours: Discussion: *Neurology (India)* **34** 155-156
- and Bhargava S 1985 Effect of medical treatment on intracranial tuberculoma A CT study; *Tuberclc* **66** 85-97
39. Tascos N A, Parr J and Gonatas N K 1982 Immunocytochemical study of the glial fibrillary acidic protein in human neoplasms of the central nervous system, *Hum Pathol* **13** 454-458
40. Vander Meulen J D M, Houthoff H J and Ebels E J 1978 Glial fibrillary acidic protein in human gliomas; *Neuropath Appl Neurobiol* **4** 177-190
41. Velaseo M E, Dahl D, Roessmann U and Gambetti P 1980 Immunohistochemical localization of glial fibrillary acidic protein in human glial neoplasmas, *Cancer* **45** 484-494
42. Zimmerman E A, Defendini R and Frantz A G 1974 Prolactin and growth hormones in patients with pituitary adenoma. A correlative study of hormone in tumour and plasma by immunoperoxidase technique radioimmunoassay, *J Clin Endocrinol Metab* **38** 577-585





**Marthanda Varma Sankaran Valiathan** (b. 24 May 1934) did F.R.C.S (1960) from England; Ch.M. (1962), University of Liverpool, UK; F.R.C.S (Cardio-vascular) (1970), Royal College of Surgeons and Physicians of Canada. He is Director, Sree Chitra Tirunal Institute of Medical Sciences and Technology, Trivandrum; Visiting Professor of Biomedical Engineering, Indian Institute of Technology, Madras.

Valiathan's contributions lie in the characterization and application of graphite benzalkonium-heparin coating of materials in cardiac surgery; development of a geochemical hypothesis in the causation of endomyocardial fibrosis and its experimental demonstration; and development of indigenous technologies for heart valve, oxygenator and cardiectomy reservoir.

Valiathan is Fellow of Indian Academy of Sciences (sometime Vice-President), National Academy of Medical Sciences (India), National Academy of Sciences (India), Indian National Academy of Engineering, American College of Cardiology, Indian Association of Cardiothoracic Surgeons, Society of Thoracic Surgeons, (USA), and Society of Cardiothoracic Surgeons of UK and Ireland; has been Member, INSA Council (1991). He is the recipient of Hunterian Professor, Royal College of Surgeons, England (1969); Kerala State Award in Science and Technology (1982); R.D. Birla Award (1986); O P. Bhasin Award (1987), Shri Dhanwantari Prize (INSA) (1991); Padma Bhushan (1990).

*Marthanda Varma Sankaran Valiathan was elected to the fellowship of the Academy in 1984*

**Address:** Director, Sree Chitra Tirunal Institute for Medical Sciences and Technology, Trivandrum 695 011; Tel. (0471) 443085; Fax: 91-471-446433, Telex: 0435-290; Telegram: CHITRAMET; Res. No.4, Kotakompally Scheme, Pulaynarakotta, Trivandrum; Tel: (0471) 448546

## **A TROPICAL HEART DISEASE REVISITED**

M S VALIATHAN

### **INTRODUCTION**

It is more than four decades since a heart muscle disease known as endomyocardial fibrosis (EMF) was reported in detail from Uganda (Davies 1948). In subsequent years it was spotted in the Ivory Coast, Nigeria, Sudan, Kerala and Brazil which share their equatorial location with Uganda. Heart muscle diseases fall into three groups—dilated, hypertrophic and restrictive—and EMF belongs to the restrictive group. It is characterised by the progressive restriction of cardiac relaxation and in late stages, of contraction. Patients generally approach the hospital for shortness of breath, swelling of the face and abdomen, palpitations and other symptoms which denote heart failure. It is seldom they are seen in the early stages of the illness. Left to medical treatment, hardly 35% of the symptomatic patients survive five years (Gupta et al. 1989). The only effective treatment for disabled patients is the surgical excision of the fibrous overgrowth within the heart and the replacement of mitral and tricuspid valves if they are engulfed in fibrosis. Surgically treated patients fare better in so far as 60% are alive at the end of 5 years (Valiathan et al. 1987).

### **PATHOLOGIC ANATOMY**

As the name implies, the hallmark of the disease is fibrosis of the endomyocardium. The fibrosis may be patchy in either ventricle, affecting the apex or the base of the anterior papillary muscle in the right ventricle or the apex or upper third of the posterior wall in the left ventricle. The preference of the fibrosis for the inflow part of the ventricles is striking. As the disease progresses, the fibrosis may involve the entire chamber, diminishing the size of the cavity and plastering the tricuspid or mitral or both valves in that process. As the valves become incompetent due to immobilisation, the right atrium dilates in particular and attains massive dimensions with the formation of thrombi. Other organs are spared except as a consequence of chronic congestive heart failure (Karthi & Sandhyamani 1985).

Histologically, the fibrosed endomyocardium shows a superficial layer of densely packed collagen with a subjacent layer of varying degrees of vascularity and myocyte degeneration. Inflammatory cells are seldom conspicuous. If the fibrosis is well established in one ventricle, the opposite ventricle, apparently normal, may show interstitial cell proliferation and mild changes of myocyte atrophy (Karthi & Gupta 1991). Whether these changes can be regarded as the early lesions of EMF is a matter of ongoing debate. At the ultrastructural level, the striking change is the marked thickening of the basement membrane of capillaries (Karthi & Valiathan 1988).

### CAUSATION

While echocardiography and surgery greatly improved the diagnosis and treatment of EMF during the last two decades, the understanding of its causation or pathogenesis made little progress in the same interval. Many theories on causation appeared over the years but none found support in confirmatory evidence. The theories variously viewed EMF as a hyperimmune variant of rheumatic heart disease, a consequence of serotonin toxicity, a sequel to viral myocarditis or a manifestation of eosinophilic heart disease. Among these theories, the eosinophilic view was vigorously propagated by Olsen who claimed that the EMF in the tropics and the Loeffler's eosinophilic heart disease of the temperate zone were manifestations of the same cardiac response to eosinophilic injury (Brockington & Olsen 1973). As the morphological picture of the heart of both conditions was very similar in the final stage, Olsen's unitarian theory became popular and tropical investigators lost no time in looking for the eosinophil factor in EMF. In the anxiety to locate an eosinophilic factor in the heart, it was forgotten that morphological similarity may be deceptive and may hold no clues to causation. Atherosclerosis progresses rapidly, if not gallops, in a donor heart from a 20-year old man following transplantation and its coronary arteries may look no different from those of a 70-year old man with advanced coronary atherosclerosis. No one would however claim that the coronary disease in both situations share an immunologic basis and a unitarian causation.

Reiteration notwithstanding, the eosinophilic theory of tropical EMF, rests on weak foundations. It does not explain the marked prevalence of the disease in the tropics or its preference for the poor. It ignores the many observations of tropical investigators that their patients have neither significant eosinophilia nor eosinophilic products in their

cardiac tissues. The unitarian theory claims that the alleged eosinophilia leading to EMF is often due to filarial infection - yet the disease is not seen in the hyperendemic areas for filariasis such as Bihar, Orissa or Tanjavur in India. While seeking a common basis for tropical EMF and Loeffler's eosinophilic heart disease, the unitarian theory glosses over the stark differences between the two conditions in age, mode of onset, symptoms and signs, multisystem involvement and eosinophilia (Valiathan & Kartha 1990). To look at the causation of tropical EMF, one is obliged to rid one eyes of the eosinophilic scales.

### A GEOCHEMICAL VIEW

In fact, what stand out in the clinico-pathological canvas of EMF are its marked preference for the tropical belt and equally strong preference for the children and adolescents among the poor. The tropical prevalence is overwhelming as 730 out of 779 cases reported during the last two decades belong to countries within  $15^{\circ}$  of the equator. Similarly, most patients—over 90% in Kerala - are very poor and they are struck by the disease in childhood or adolescence. Any hypothesis is obliged to account for these two features of EMF besides suggesting a basis for its cardiac selectivity.

As the tropical belt shares latasolic soils with abundant minerals, an analysis was made several years ago to determine whether the elements in the soil might have a connection with the genesis of EMF. The initial report suggested that the heart tissue samples in patients contained less magnesium and more thorium than the control samples (Valiathan et al. 1986). A subsequent study revealed that the concentration of cerium in the patient's heart tissues was much higher than that of thorium (Valiathan et al. 1989). As cerium constitutes 30% by weight of monazite and thorium accounts for only 7%, these findings suggested a possible role for monazite as their source. Moreover the levels of magnesium and cerium raised the possibility that magnesium deficiency might have enhanced the level of cerium and provided binding sites for the toxic metal in the cardiac tissues. As magnesium is essential for a number of reversible reactions involving energy transduction, its replacement by a toxic element like cerium would make the reactions irreversible and pave the way for cellular dysfunction and death. These suggestions shifted the causation of EMF to the nontraditional context of geochemistry.

The geochemical hypothesis raised a number of questions. They are discussed below in juxtaposition with the answers from experimental studies.

### CAN CERIUM REPLACE MAGNESIUM ?

Cerium and magnesium occupy positions which are far apart in the periodic table. They differ markedly in atomic number, ionic radius, charge, electronic configuration and other physico-chemical characteristics. Therefore the theoretical question whether a heavy element like cerium can replace a light element like magnesium is appropriate. In an exhaustive discussion on the replacement or substitution of elements, Jacobson and Turner pointed out that the charge-radius ratio out-weighed individual physico-chemical properties in determining the exchange or substitution of elements (Jacobson & Turner 1980). Therefore it is of interest that the charge-radius ratios of magnesium and cerium are 3.03 and 2.90 respectively and the closeness in ratios may outweigh their differences in the present context of possible substitution. Moreover direct evidence for the substitution of magnesium by cerium was obtained by Shivakumar who studied the binding of creatine kinase with cibacron blue F-GA, its substrate. He showed that the binding which requires magnesium as a co-factor could occur just as readily with cerium as a co-factor (Shivakumar et al. 1989).

### ARE THE LOWER CONCENTRATION OF MAGNESIUM AND HIGHER LEVEL OF CERIUM FUNCTIONALLY COUPLED ?

The question arises whether the levels of magnesium and cerium as observed in the cardiac samples of EMF are inter related or whether they are independent of each other. This question needs to be considered in the context of data which show that the deficiency of magnesium enhances the concentration of lead in gestant rats and puppies and that it may account for the higher levels of tissue aluminium in neuro-muscular conditions (Cerklewski 1983, Gajdusek & Jalazar 1981). In other words, magnesium deficiency is known to play a synergistic role in enhancing the concentration, and possible toxicity, of trace metals. Nair et al. demonstrated in a tissue culture model of a tuber crop-*Coleus parviflorus* - that the deficiency of magnesium did enhance the concentration of cerium (Nair et al. 1989). In another study Shivakumar et al. showed that magnesium deficiency had a similar effect in raising the cardiac level of cerium in primates (Shivakumar et al.) (unpublished observation). These

findings suggest that the observed levels of magnesium and cerium in the myocardial samples of EMF are functionally coupled.

### DO THE LEVELS OF MAGNESIUM AND CERIUM CORRELATE WITH THE CLINICAL AND EPIDEMIOLOGICAL FEATURES OF EMF ?

Magnesium deficiency commonly accompanies malnutrition which is marked in EMF patients. This has been reported from Africa as well as Kerala (Caddel 1969, Eapen et al. 1991). Magnesium deficiency occurs due to poor food intake and it is aggravated by diarrheal diseases which inhibit the absorption of magnesium from the gut. Given the higher growth needs of magnesium in children and adolescents, magnesium deficiency will make them more vulnerable to the disease.

However magnesium deficiency alone causes multi-system damage (Heggtveit et al. 1964) which is absent in EMF where its role would seem to be synergistic. Probably magnesium deficiency enhances the absorption of cerium from the gut and provides binding sites in the cardiac tissues for cerium.

If poverty and malnutrition operate through magnesium deficiency, the equatorial incidence of EMF acts through cerium which is the major elemental component of monazite. And monazite is abundant in the latasolic soil of the equatorial belt.

### IS CERIUM BIOLOGICALLY ACTIVE AT NANOGRAM LEVELS ?

Even though the cerium concentration in the cardiac tissues of EMF is significantly higher than that of controls, the fact remains that the levels are measured in nanograms. Is this functionally significant ? Is it, as a critic might observe, an instance of data unearthed by powerful analytical instruments searching for significance even as a cause might look around for a disease ? These doubts have been set at rest by a recent study which clearly showed that cerium has a stimulatory effect on collagen synthesis at nanogram level whereas it tends to be inhibitory at microgram level (Shivakumar et al.) (in press). Therefore a role for cerium in the fibrogenesis of EMF can no longer be ruled out.

## DOES THE REPRODUCTION OF MAGNESIUM AND CERIUM LEVELS SEEN IN EMF INDUCE SIMILAR MORPHOLOGICAL CHANGES IN AN EXPERIMENTAL MODEL ?

Over a two year period, a number of experiments have been carried out to develop an experimental model of EMF in rat. The supply of magnesium deficient feed, standardising the desirable level of magnesium, the stage, doze and route of administration of cerium and several other experimental conditions had to be established and the data obtained in a statistically acceptable form during the study. At the present time one subset of animals who had 'acute on chronic' magnesium deficiency and cerium supply in drinking water have shown marked endocardial lesions. These findings are not significant statistically and further experiments are under way (Karthi et al. 1991). As rats differ from the humans in rapidly eliminating cerium from the body it may also turn out that a large animal is a more appropriate model for the study.

## CONCLUSION

The stakes in creating an experimental model are high not only for the geochemical hypothesis : they are equally high for the potential patients. This is because the success of the model will open the prospect that the disease may be prevented or turned round by magnesium supplementation. No more than a hope at this time, its elaboration must await another day.

## References

1. Brockington I F and Olsen E G J 1973 Löffler's endocarditis and Davies endomyocardial fibrosis; *Am. Heart J* 85 308-322
2. Cerklewski F L 1983 Influence of maternal magnesium deficiency on tissue lead content of rats; *J. Nutr.* 113 1443-1447
3. Caddel J L 1969 Magnesium deficiency in protein-calorie malnutrition : A follow up study; *Ann. N Y Acad. Sci.* 162 874-890
4. Davies J N P 1948 Endomyocardial fibrosis in Africans; *East Afr. Med. J.* 25 10-14
5. Eapen J T, Rajashree S, Radhakumari C and Nair R R 1991 Serum magnesium levels in school children of two socio-economic groups in Kerala; *Proc. Int. Symp. Endomyocardial Fibrosis* (in press)

6. Gajdusek D C and Jalazar A M 1981 Amyotrophic lateral sclerosis and Parkinsonism syndromes in high incidence among the Auyu and Jakai people of West New Guinea; *Neurol* 32 107-126
7. Gupta P N, Valiathan M S, Balakrishnan K G, Kartha C C and Ghosh M K 1989 Clinical course of endomyocardial fibrosis; *Br. Heart J.* 62 450-454
8. Heggveit H A, Hereman L and Mishra R K 1964 Cardiac necrosis and calcification in experimental magnesium deficiency: a light and EM study; *Am J Path* 45 757-782
9. Jacobson K B and Turner J E 1980 The interaction of cadmium and certain other metal ions with proteins and nucleic acids; *Toxicol* 16 1-37
10. Kartha C C and Sandhyamani S 1985 An autopsy study of tropical endomyocardial fibrosis in Kerala; *Indian J Med. Res* 82 439-446  
 — — and Valiathan M S 1988 Cardiac ultra-structure in tropical endomyocardial fibrosis; *Indian J Med Res.* 87 275-282  
 — — and Gupta P N 1991 Pathologic spectrum and possible pathogenesis of endomyocardial fibrosis; *Proc Int Symp Endomyocardial Fibrosis* (in press)
11. Nair R R, Gupta P N, Valiathan M S, Kartha C C, Eapen J T and Nair N G 1989 Enhanced cerium concentration in magnesium deficient plants; *Curr Sci* 58 696-697
12. Shivakumar K, Appukuttan P S and Kartha C C 1989 Interactions of thorium and cerium with adenine nucleotides; *Biochem Int.* 19 845-853  
 — —, Nair R R and Valiathan M S Paradoxical effect of cerium on collagen synthesis in cardiac fibroblasts; *J Mol Cell Cardiol* (in press)  
 — —, Eapen J T, Ramesh V and Valiathan M S Magnesium deficiency promotes the concentration of cerium in the cardiac tissue in the non-human primate, *Macaca radiata*; (unpublished observations)
13. Valiathan M S, Kartha C C, Panday V K, Dang H S and Sunta C M 1986 A geochemical basis for endomyocardial fibrosis; *Cardiovasc Res* 20 679-683  
 — —, Balakrishnan K G, Sankarkumar R and Kartha C C 1987 Surgical treatment of endomyocardial fibrosis; *Ann. Thorac. Surgs.* 43 68-73  
 — —, Kartha C C, Eapen J T, Dang H S and Sunta C M 1989 A geochemical basis for endomyocardial fibrosis; *Cardiovasc Res.* 23 647-648  
 — —, and Kartha C C 1990 Endomyocardial fibrosis: the possible connection with myocardial levels of magnesium and cerium; *Int. J. Cardiol.* 28 1-5



**Amulya Kumar N Reddy** (*b.* 21 October 1930) did Ph.D. (1958) from Imperial College of Science and Technology, University of London, UK. He is Chairman, Department of Management Studies, and Professor, ASTRA (Centre for the Application of Science & Technology to Rural Areas) and Department of Inorganic & Physical Chemistry, Indian Institute of Science, Bangalore, India.

Reddy has made major contributions to the fields of Global energy strategies, energy problems of developing countries, rural energy consumption patterns, rural energy centres, biogas technology, village ecosystems, biomass potential and utilization, science policy for developing countries, technologies for rural development alternative/appropriate technologies. He has also undertaken electron diffraction studies of electrodeposits, structure and growth of electrodeposits, electrodeposition, corrosion, ellipsometric studies of anodic films, thin layer electrochemistry, ionexchange membranes and nerve conduction. He has Co-authored 'Modern Electrochemistry' (Plenum Press, New York, 1970); 'Energy for a Sustainable world' (Wiley- Eastern); authored 'Science, Technology and the Environment— A Reappraisal' (United Nations Environment Programme, 1978); and edited 'Rural Technology' (Indian Academy of Sciences, Bangalore, 1980).

Reddy is Fellow of Indian Academy of Sciences, Bangalore. He is the recipient of Rathindra Puraskar (Viswabharati University, Shantiniketan) (1982) and B.D.Tilak Lecture Award (Indian National Science Academy) (1985)

**Address:** Chairman, Department of Industrial Management, Indian Institute of Science, Bangalore 560 012

## BIOMASS AND RURAL DEVELOPMENT

AMULYA KUMAR N REDDY

### INTRODUCTION

Any discussion of rural development is based, either consciously or unconsciously, explicitly or implicitly, on a viewpoint on rural development. It is best, therefore, to preface the present exposition with the viewpoint on rural development that underlies it.

A major handicap in such a discussion is the shibboleth that development can be equated with growth, and that as a corollary, growth maximization should be the objective of development.

The experience of developing countries points in another direction. Growth maximization has only led to greater polarization into "dual" societies because the benefits of growth have been skewed in favour of the minority elites and against the poverty-stricken majorities. The problem is that the content and structure of growth are as important as its magnitude. What goods and services are produced are as important as the quantities in which they are produced. Economic growth, therefore, is a necessary condition for development, but *not a sufficient condition*. Development must be an economically-efficient process of economic growth directed towards :

- (a) *equity* through the satisfaction of basic human needs, *starting from the needs of the neediest*
- (b) *environmental soundness* to make development sustainable over the long run
- (c) *self-reliance* to ensure a participatory development.\*

Thus, rural development should be viewed as a needs-oriented, environmentally sound, self-reliant process of economic growth. Since rural development cannot be isolated from urban development, it is

---

This lecture is by way of a personal tribute to the social sense and enlightened philanthropy which Prof. Tilak has shown through his attempt to conscientize Indian scientists and engineers regarding the exciting and inspiring tasks of rural development.

obvious that the two processes must proceed symbiotically, and that urban growth must not be parasitic and at the expense of rural areas.

Biomass and rural development are inter-connected in many ways. Biomass production is essential for meeting the development needs of food (grains, vegetables, fruits), fibre for clothing, forest (non-fuel) products for shelter (lumber), clothing (rayon), health (medicines), and communication (paper). Biomass production needs energy inputs, the provision of which are vital developmental tasks. And, biomass is (and can continue to be) a major fuel source in the energy system.

### ENERGY FOR BIOMASS PRODUCTION

Biomass production is primarily a matter of agriculture (including silviculture). Traditional agriculture uses only animate energy sources (human beings and draught animals); it also depends overwhelmingly on locally available inputs. All this makes it self-reliant, but a price may have to be paid—its productivity may be too low to meet the needs of the growing population. Agriculture may have to be modernized.

*Agricultural Modernization* can be looked on as a process of replacing traditional inputs (implements, seeds, organic fertilizers, animate energy sources) with "modern" inputs (e.g., fertilizers, pesticides, herbicides, etc.), mechanical equipment (e.g., seed drills, sprayers, tractors, etc.), and "modern" energy sources (electricity, diesel, etc.)

Consider as an example four technologies of rain-fed rice production which is a crop of major importance: the traditional technology of rice production, and three variants of modern technology.

Traditional rain-fed rice technology does not use hybrid seeds, chemical fertilizer, pesticides, or herbicides. It does not use either transport vehicles or draught-power sources running on oil-derivatives. Instead, it relies wholly upon: (i) draught animals for land preparation, (ii) traditional seed varieties and organic fertilizer, (iii) manual transplanting, harvesting, threshing and winnowing.

*Variant 1* of modern rice technology differs from traditional in that it uses modern biological-chemical inputs (hybrid seeds + fertilizer + insecticide + herbicide), and mechanically driven transport vehicles.

*Variant 2* of modern rice technology differs from Variant 1 by using tractors for ploughing whilst still retaining draught animals for harrowing.

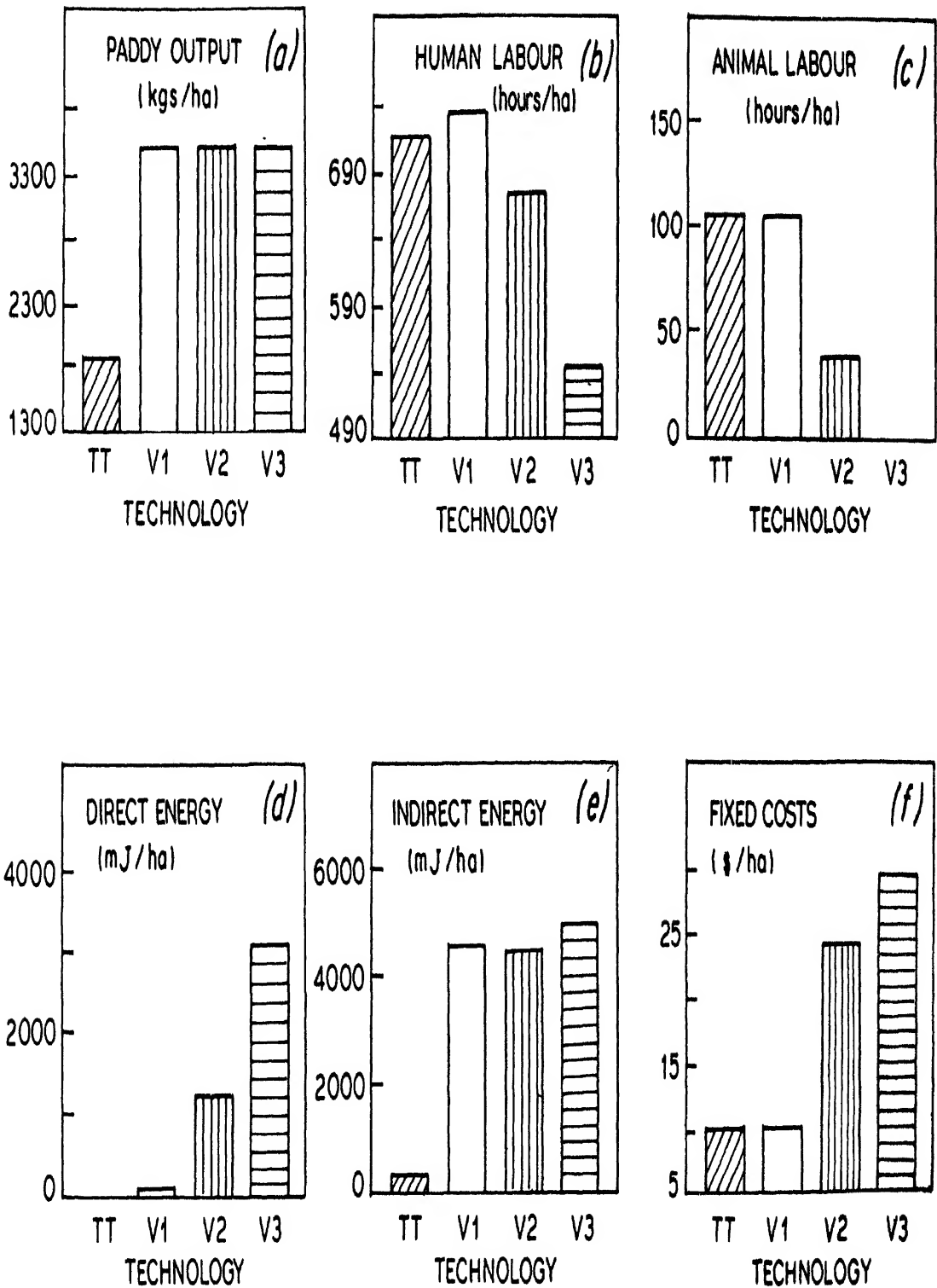


FIG. 1A A comparison of the productivity of various technologies

*Variant 3* of modern rice technology utilizes power threshers and replaces the tractors of *Variant 2* with power tillers for ploughing and harrowing.

Thus far, agricultural crop production has only been analysed with a "black-box" approach, in which the inputs (land, labour, fertilizer, water, energy, etc.) to and outputs (production) from the black-box are counted. Such an approach does not facilitate an understanding of what is happening inside the "black-box". Even when energy is disaggregated, it is done sectorwise.

What is needed is the disaggregation of agricultural crop production according to operations which are the *end-uses* of agriculture. Such a disaggregated end-use approach enables the computation *operation by operation* of the human labour, animal labour, direct inanimate energy, indirect inanimate energy, fixed capital costs and paddy output for the four technologies considered above. It constitutes the "unpackaging" of the technological "black boxes" for rice production, and makes it possible to understand the labour (including animal labour), energy (direct and indirect) and costs associated with the operations of each technology. Finally, it permits an illustrative panoramic view of the overall main "first-approximation" implications of the three technologies of agricultural modernization as the higher-level outcome of aggregating the disaggregated data at the lower operations level, and thereby provides the "raw material" and data-base for the choice of technologies.

A comparison of the productivity of the various technologies shows that the "doubling" in paddy output (figure 1A) achieved by "modernization" comes primarily through the adoption of the improved seeds, fertilizer and pesticides (*Variant 1*)—the mechanization of ploughing (*Variant 2*) and other operations (*Variant 3*) do not produce further improvement to any significant extent.

The per hectare human labour requirement (figure 1B) actually increases with the *Variant 1* of modernization because of the increased yield. Mechanized ploughing (*Variant 2*) with its greater efficiency only produces a slight decrease in labour requirements because of a decrease of the labour time required for ploughing (because of the significant increase in the ploughing rate). In contrast, *Variant 3* has a major impact on human labour requirements, because it includes the replacement of manual threshing with power threshing. It not only drastically reduces the labour requirements for threshing but also those for the traditionally subsequent

winnowing operation which is made redundant—all told, an approximately 25% reduction in the labour is traditionally required.

Variant 1 depends upon draught animals to the same extent as traditional technology and therefore shows the same animal labour requirements (figure 1C). There is a fall in the animal labour requirements only in Variant 2 where the ploughing operation is mechanized, but the requirement is "non-zero" because the harrowing operation is still carried out with animal power. Variant 3 of modern technology deals the *coup de grace* to draught animal power which is replaced not only in the ploughing but also in the harrowing operations of land preparation. As a result of draught animals being completely replaced with machines in Variant 3, the animal labour requirements become zero.

The direct energy requirements (figure 1D) are virtually "zero" in both traditional and Variant 1 technologies because neither of them involves any mechanization. But, Variant 2 shows a significant requirement because of the mechanization of the ploughing operation, and there is a dramatic additional increase in direct energy use in Variant 3 because of the further increase in the extent of mechanization.

In the case of indirect energy (figure 1E) which is the energy used in the manufacture of the inputs, all the variants of modern technology show a major increase in requirement (approximately 14 times compared with traditional technology) because of the off-farm consumption of indirect energy for the production of improved seeds, fertilizers, insecticides and pesticides. It follows that Variants 1, 2 and 3 are associated with major escalations in variable costs on account of the purchase of inputs that are intensive in indirect energy. They also involve dependence on external agencies which have to be established for the supply of seeds, fertilizers, equipment, fuels, etc., manufactured in urban industries servicing the agricultural sector.

Considering direct and indirect energy together, it is easy to see why modern agriculture is such a major energy sink compared to traditional agriculture.

Finally, figure 1F shows the variation of the fixed costs with the technologies—the introduction of mechanical equipment with Variants 2 and 3 leads to sharp increases (by factors of 2.5 and 6.3 compared to traditional technology) in these costs which do not show much of a change in going from traditional to Variant 1 technologies.

Thus, the "mechanized" technologies of rice production, viz., Variant 2 and Variant 3 (minus its power thresher) involving "mechanized" land preparation and transport do not increase the paddy output (unless by speeding-up operations they make multi-cropping possible). They arise from an impetus to replace draught animals with tractors and power tillers in order to overcome the constraint of draught animal power, of which there may be a shortages because of: (i) the scarcity of pasture land and fodder, or (ii) the unpleasant arduousness and drudgery of using animal power. They also eliminate the necessity of managing huge numbers of cattle and cattle-herds that would be required by large holdings because this may not be as economical as in the case of small holdings.

Variant 3 of modern technology involving the replacement of manual threshing with a power thresher leads to a significant reduction in labour requirements which must be seen against the reduction in losses in grain incurred during traditional threshing. This raises two questions:

- (a) Are such labour-saving techniques introduced in pace with the rising demand for labour in urban and/or rural off-farm industry and therefore with the decreasing availability of agricultural labour? or
- (b) Are they brought in because of price distortions embodied in or backed by associated subsidies, incentives, etc.?

Though the duration of a crop is largely governed by crop physiology, the overall time taken for the completion of all the crop operations can be much greater due to the time spent on land preparation, harvesting, threshing, etc. If, therefore, these operations are speeded up, and a short-duration variety is chosen, it may become feasible to go through another crop cycle within the same agricultural year. In such a case, the energy, human labour (both human and animal), capital, costs and output would have to be multiplied by as many crops as there are in a year. For instance, if variant 3 of modern technology permits *two* rice crops per year, and traditional technology only *one*, the annual labour requirement is 1090 hr/ha which is more than the 717 hr/ha of traditional technology. Hence, the comparison of technologies must be done on a *per year* basis, and not simply on a *per crop* basis.

The substitution of the traditional winnowing operation by power threshers results in a selective reduction in employment of women with all the attendant impacts on development, in general, and their incomes and

status in particular. Thus, agricultural modernization can not only lead to a displacement of human labour, but there can be a gender bias in this displacement of labour by machines with women bearing the brunt of the process.

The analysis of rice production technologies has shown that agricultural modernization does not consist of a unique inflexible package which must either be accepted in *toto* or not at all. It appears that modernization could be accomplished in several stages, for example:

- Stage 1 : Green revolution without either tractors or machines,
- Stage 2 : Green revolution with tractors and power tillers,
- Stage 3 : Green revolution of the industrialized-country type in which virtually all the operations including harvesting, threshing, etc., are mechanized.

The discussion above has focussed a great deal of attention on energy consumption in agricultural biomass production. But, energy cannot be the sole determinant in the choice of agricultural technologies. It is only one among several other factors in the domain of decision-making, such as agricultural productivity, employment generation, investment costs and environmental degradation. In fact, the basic developmental issues to be considered in agricultural modernization include: (1) equity, (2) employment, (3) ecological impacts, and (4) energy consumption.

Nevertheless, an important issue concerns the source(s) of the energy inputs for raising agricultural production. The crucial question is whether biomass sources of energy can meet the requirement of agricultural production.

### BIOMASS FOR ENERGY PRODUCTION

The main biomass sources of energy are fuelwood, agricultural wastes, and animal wastes.

The term fuelwood covers logs, branches, twigs, roots, bark, etc., from trees, bushes, shrubs, etc., located not only in forests and woods but also along the sides of roads, tanks, fields, etc. Fuelwood is used primarily in: (i) the domestic sector (for cooking and water heating), and (ii) rural industries (brick-making, jaggery, pottery, black-smithy, etc.).

Agricultural wastes are wastes in the form of stalks, leaves, husk, etc., left after saving the grain, fibre, etc., parts of the crops. Of these, only



bagasse and rice-husk are used in significant quantities as fuels, the other wastes are consumed as fodder or deployed as green manure.

Animal wastes are mainly cattle dung which is used as organic fertilizer and in the form of dried cakes as fuel.

Biomass has played and still plays a crucial role in India's national, urban and rural patterns of energy consumption.

Biomass accounts for about 55% of the total energy used in the country (9000 PJ in 1978). The total biomass consumption in this year was stated to be 295 MT with households accounting for 245 MT and industries, 50 MT. Within the domestic sector, the breakup was as follows: fuelwood, 132 MT; agricultural wastes, 41 MT; and animal wastes, 72 MT. A number of recent studies have shown that even in modern cities, fuelwood consumption is not insignificant. For example, in the city of Bangalore, the daily fuelwood consumption was about  $1200 \pm 50$  tonnes per day corresponding to about 150 kg/capita/year.

Since the role biomass in the urban and national settings has been discussed elsewhere, the focus here will be on biomass in rural energy consumption patterns. Even taking into account animate energy—the energy of human beings and draught animals—in computing the total energy consumption, biomass can contribute as much as about 90% of the energy used in a village.

Fuelwood is the main fuel, because on the one hand, dung cakes are, in general, used as fuel only where fuelwood is not available within a convenient distance, and on the other hand, agro-wastes are a major source of fodder. The fuelwood consumption is about 500 kg/capita/year which corresponds to about 10 kg/household/day.

Fuelwood, agro-wastes and animal wastes are referred to as non-commercial energy because they are gathered at "zero" private cost, but they are rapidly becoming commercialized. Fuelwood is gathered by women and children in the form of twigs. It is used for cooking (about 80%), water-heating and industry. Fuelwood stoves are multi-pot mud-stoves with about 8% efficiency (which becomes about 16% taking water evaporation as useful work).

Biomass has an important role to play in meeting village energy needs. In analysing this role, priority must be assigned to cooking which is the largest and most important energy consumer. There are many options for meeting cooking energy needs: (i) the present pattern of consumption

(including the extremely low efficiency) could be preserved and efforts concentrated on increasing fuelwood supplies—this is the *forestry* option; (ii) The end-use efficiencies in cooking could be improved—this is the improved high-efficiency fuelwood *stoves* option; (iii) A switch could be made from fuelwood to charcoal or biogas or producer gas—this is the *alternative fuels* option; (iv) There could be a reduction of demand with conservation of supply—this is the *demand-decrease-supply-step-up synergism* option.

The first general issue in making these choices is that of *increasing supply vs managing demand*. Just as it is invariably cheaper and quicker to save a kilowatt than to generate a kilowatt, the biomass version is that it is cheaper and quicker to save fuelwood than to grow fuelwood. However, saving alone may not be enough, and it may be imperative to exploit the demand-decrease-supply-step-up synergism.

The second issue concerns the question of *solid vs gaseous cooking fuels* noting that gaseous fuels are not only easier to light and extinguish but they facilitate quick increases of power output (for boiling) and decreases of power output (for simmering).

The third issue is whether the present practice of *solid fuel for the poor and gaseous fuel for the rich* should be permanent or temporary, i.e., whether a dual cooking-fuel energy system should consolidate a dual society or whether solid fuel should be transitional fuel to the ultimate gaseous cooking fuel(s).

The fourth issue is that of *centralized large-scale generation plus piped distribution of not-easily-liquefiable gas vs decentralized small-scale generation plus consumption*, i.e., the economies of scale vs diseconomies of organization and management.

In considering these issues, note must be taken of the recent exciting developments consisting of *doubling* of stove efficiencies and dissemination of improved stoves. Unfortunately, the poor who now use 'zero-cost' stoves cannot afford the improved stoves even though they only cost Rs. 50–100—hence, these stove programmes are based on subsidies.

Stoves R & D has also led to several lessons, such as the importance of: (i) "market surveys" in the form of studies of actual cooking practices and needs; (ii) basic science, because stoves are not a trivial R & D

problem; (iii) "test-marketing" in the form of user trials; and (iv) "field testing" with statistically significant samples.

Experience has also been gained with biogas plants. Firstly, as was predicted a decade ago, family-size biogas plants have extremely limited possibilities with regard to diffusion because they are not cost-effective without massive subsidies and because only a small percentage of affluent farmers have the cattle resources to sustain the inputs to these plants. Secondly, recent experiments with community biogas plants have stressed the importance of: (i) involvement of the community, (ii) a "utility" approach, and (iii) overcoming the (dung) resource constraint associated with the low body weight of cattle in drought areas through the integration of biogas programmes with dairy development and fodder production programmes.

Lighting is an end-use which does not account for a large fraction of the total energy budget of a village, but it has a dramatic impact on the quality of life. The bulk of rural households depends on kerosene for illumination—on the average, over 80% of rural homes are unelectrified even in electrified villages. This is the case even though the kerosene lamp has a luminous efficiency which in order of magnitude is lower than the incandescent bulb which is the least efficient of the family of electric illumination devices that includes the fluorescent tube. In addition, the use of kerosene for lighting involves several socioeconomic penalties that only confirm that electricity is the most appropriate energy carrier for illumination.

There are two options for obtaining the electricity to illuminate village homes: the electricity generated centrally at distant power plants can be transmitted over grids and distributed to villages—the well-known rural electrification programme—or it can be generated in a decentralized way from locally sited generating sets. There are break-even distances from the grid within which it is more economical to extend the grid to the village, and above which decentralized local generation is more sensible. Of course, the break-even distance depends upon the fuel for the local generator, and the obvious fuels are diesel, biogas and wood/producer gas. Unfortunately, a rigorous comparison between these three fuels has not been made.

Heating bath water is also an important end-use in parts of the country where the nights are cool and, therefore, it is the practice for people to use hot water for bathing. The hot water is obtained by burning

wood fuel under cauldrons of water to raise its temperature to about 50-60°C.

The technological problem of heating water is quite similar to that of cooking from which, however, it differs in two important ways; (i) there is no need for the heated water to reach temperatures near the boiling point, and (ii) the volumes of water to be heated are at least an order of magnitude higher. The larger volumes to be heated, and therefore the larger energy requirements, militate against the use of biogas which is a much scarcer fuel than was once thought to be the case. The lower temperatures bring solar water-heating within the range of acceptability provided that it is done on a community-scale with hot-water distribution at a few selected public outlets so that the costs of piping hot water to individual homes is avoided. Such a community approach to solar water-heating may well involve a host of organizational problems, but the fact is that it has not yet been tried out.

In this context, it appears that fuelwood-fired water-heating stoves are the most sensible option. It may well be the case that the efficiencies of traditional *water-heating* stoves are higher than those of traditional fuelwood *cook* stoves because of the well-known efficiencies of scale of thermal reactors. Despite this, there is urgent need to draw upon the research and development experience with wood-fired cookstoves and develop improved fuel-efficient water-heating woodstoves.

In most parts of the country, agriculture is neither wholly traditional nor is it completely modernised; it is in fact passing through a transitional phase. It is characterised by dependence on draught animals and human energy and on traditional implements, but it also uses chemical fertilizer in irrigated lands as well as high-yielding varieties (HYV) on some of the farms.

Since crop productivity is highly sensitive to energy, nutrients and water, the provision of these inputs—the ingredients of agricultural modernization—is an important developmental task. Thus, agriculture needs direct energy for agricultural operations and for pumping irrigation water, and indirect energy as fertilizers.

Though draught animal power is a self-reliant, local, non-renewable source of energy, it leads to certain limitations on crop output. Thus, it has been shown that the existing draught power is inadequate for achieving timely sowing in its dry lands. This problem arises because the "ploughing

window"—the period after the ploughing rains during which the soil is moist enough for the given power sources, say draught animals, to carry out the ploughing operation—is too narrow for draught animals to complete the ploughing job on the dry lands. This situation may warrant the deployment of tractors and/or power tillers (not necessarily privately owned, but for instance hired out) for the timely completion of the land preparation and sowing operations.

The inadequacy of power delivered by draught animals can also prove an obstacle to increasing the cropping intensity. When only one cycle of crop production is considered, it can be argued that mechanization is quite unnecessary unless there is a shortage of labour. Hence, in single crop areas it may be necessary to improve the efficiency of animal energy use.

If, however, the objective is to increase the cropping intensity by growing more than one crop per year, then it becomes important to consider the total time for the completion of a crop cycle. Like the crop output, this cycle time too is essentially determined by the biological factor which in this case is the rate of growth of the crop variety. When, however, short-duration crop varieties are adopted, the possibility of double, and particularly triple cropping may depend upon reducing the time taken for land preparation (ploughing, harrowing and interculture), harvesting, threshing and other operations. In such a situation, mechanization of these agricultural operations may become necessary to double or treble the cropping intensity and thereby increase crop production/ha as well as employment to process the increased output.

Another aspect of the problem of replacing draught animals with machines derives from the low efficiency with which draught animals convert their biomass fodder input into work output. If the calorific value of the fodder is considered, it turns out that the energy equivalent of the biomass input is converted into animate energy output with efficiencies as low as 4 to 7%. If, on the other hand, this same biomass input had been converted into inanimate energy through devices such as woodgas engines, the efficiency would have been much higher. In many villages, it can turn out that draught animals delivering mobile and stationary power consume every year almost as much biomass (in the form of dry fodder) as the entire amount used for cooking, water-heating and process heating. If, however, the animal power is replaced by say woodgas engines, then there

is a saving of biomass to the extent of about 90% of the amount consumed by the animals.

Further, since crop residues cannot provide the fodder required for the maintenance of the draught animals, land which could otherwise have been used for growing cereals and pulses has to be diverted for the purpose to produce fodder crops or set aside for grazing. In many villages, the area devoted to fodder production is about one-third the area devoted to crops. In comparison, less than 10% of the land now used for fodder would be required to grow the wood to run woodgas engines and deliver power to accomplish the same job as the draught animals. Thus, the inefficiency of draught animals as energy-conversion "devices" also implies a low efficiency of land utilization. This is why in regions such as the Punjab in North India, where land is extremely precious, animals are being replaced by machines.

The upshot of this discussion is that, at some stage of the development process, it may become necessary to replace animal-powered agriculture with mechanical system powered by sources such as wood gas.

Energy for pumping water for irrigation is the other major energy requirement in the agricultural sector. Presently, it comes predominantly from electricity from the grid even though the transmission and distribution costs of grid electricity are highly subsidised. In addition, diesel irrigation pumpsets are also used as substitutes or as back-up equipment. It would be advantageous, however, to run pumpsets with local, self-reliant renewable sources of energy. The power requirement for lift-irrigation is characterised by peak and lean periods and the pumps are scattered over the farm-lands of the village. Wood (or charcoal) or electricity (generated in the village or nearby) are the obvious options for the energy carrier which has to be transported/transmitted to the dispersed pumpsets.

Woodgas-driven pumpsets seems to be a particularly attractive option. The wood required, assuming a 5 hP engine working for 300 hr/year, comes to 1.3 tonnes/year/pumpset. In other words, at 6 tonnes/ha/year, 0.2 ha of energy forest are required to supply the wood requirements of each pumpset in a renewable, environmentally sound and self-reliant way that also creates employment locally.

## THE FUTURE ROLE OF BIOMASS IN RURAL DEVELOPMENT

Biomass-based energy sources and devices, therefore, are obvious candidates for the energy-utilizing activities in the developmental scenario for villages. However, before they become valid options and are included into this scenario, it must be shown quantitatively that at least the present energy needs of the village could be met from local renewable biomass resource.

Biomass can be utilized in the form of fuelwood and crop residues for heating (cooking in households and coffee-shops, heating bath water, and jaggery making), generating stationary shaft power for irrigation water and for crushing sugarcane and oilseeds, and providing mobile power (for agriculture and transport). Also, biomass in the form of cattle waste (which is now being used only for manure) can be utilized for producing biogas without losing the manurial value of the dung. This biogas in turn can run biogas engines pumping domestic water and for generating electricity via gensets for lighting. Of course, there is no sanctity to this particular pattern of usage of biomass, and innumerable other schemes can be imagined. In fact, many of them are being implemented.

Calculations of the quantitative requirements of biomass in the form of fuelwood/agro-wastes and of dung reveal several points. Firstly, it appears that locally available biomass resources provide a feasible basis for the energy component of village development. Secondly, if the new biomass schemes are based on fuel-efficient technologies, then they lead to a definite conservation of biomass resources. Thirdly, the shift to efficient new technologies results in biomass—which has been the dominant fuel in villages—assuming an even greater role in advancing current developmental needs. Finally, this enhanced role becomes possible because of the significant savings of biomass resulting from these new technologies and the diversion of these savings to new uses.

Thus, a greater role for biomass energy resources can be achieved if and only if efficient new technologies are adopted. Had traditional technologies been retained, and had the conservation of biomass resources not been there, then the supply of biomass production would have had to be increased inordinately. And these increases may stretch village ecosystems beyond their resource "limits". This demonstrates the limitations of a purely "supply approach" of persevering with inefficient traditional technologies and trying to bridge the demand-supply gap with increases in the supply of energy.

It must be stressed that the above observations must, in general, be restricted to *present* energy needs. The complexion of the problem may change drastically when the focus of discussion changes to *future* energy needs. Attempts have, therefore, been made to examine the role of biomass against a background of village energy needs in an imaginary future.

This exercise has been carried out by: (i) estimating the useful energy corresponding to the present consumption of final energy in villages, (ii) escalating this useful energy to reflect an improvement in the standard and quality of life, (iii) assuming new technologies with significant increase in the efficiency of energy use, and (iv) computing the corresponding requirements of (final) energy. The future requirements of useful energy for villages can be based on an "utopian" scenario involving major increases in the requirements of cooking, heating bath water, process heating, lighting, lift-irrigation, mobile power for farm equipment and for transport, and shaft power in industry.

The conclusion that emerges is clear: while a purely conservation approach may be adequate for present needs, it cannot sustain future requirements *unless there is also an increase in supplies*. Apart from using trees efficiently, it is also vital there is an increase in the productivity with which trees are grown. If biomass is going to be increasingly used for meeting the energy needs of the present *and the future*, then it is imperative that biomass production must be increased which means that in in most parts of India—the biomass productivity must be increased.

It is widely believed that such a biomass productivity increase necessarily requires a change in the pattern of energy sources for agriculture (including silviculture) involving a shift to inanimate sources for various agricultural operations and to indirect energy in the form of seeds, fertilizers, pesticides, etc. Even if this shift is only partial, it means that an increase in biomass productivity hinges on energy inputs into agriculture. Thus, through the working out of its own logic, a biomass-for-energy-production strategy necessarily leads back to an energy-for-biomass-production strategy.

### THE BIOMASS-ENERGY FEEDBACK LOOP

The fundamental probable of *biomass and rural development* can be stated thus. Rural development necessarily means increasing production



biomass for food, fibre and forest products. But this task may be achieved with traditional biomass-production technologies which are wholly inadequate. It is essential to go in for improved biomass-production agricultural (including silvicultural), *tree-based* and/or *crop-based*, technologies.

But modern agricultural technologies in their present form give rise to several important questions pertaining to the long-term sustainability of monocultures of various species and to the ecological consequences of the chemical fertilizers and pesticides used today. Hopefully, these ecological problems will be solved through recent advances in molecular biology and genetic engineering.

Moreover, modern technologies for increasing biomass production require inputs of direct and indirect energy, and the magnitude of these energy inputs is a function of the particular variants of modern technology that are adopted. Here, it must be noted that there are variants of modern agricultural technology which when introduced prematurely impede need-oriented, environmentally sound self-reliant development. Only these development-oriented variants of modern agricultural technologies for increasing biomass production must be chosen which are compatible with other developmental objectives such as employment generation.

Thus, *energy for biomass production* means increasing agricultural production through *selective* use of inanimate energy inputs for selected operations. But, are these energy inputs available? Currently, modern agriculture depends largely on electricity for stationary equipment and oil for draught and shaft power. However, sustainable agriculture requires energy sources consistent with the developmental objectives of environmental soundness and self-reliance. What are required are sustainably produced locally available or produceable energy sources.

If the solution is *biomass for energy production*, then is there enough biomass to spare for diverting energy to the agricultural sector when there are already competing demands from the domestic and transport sectors? This question arises because, apart from increasing agricultural production, development also means increasing the provision of energy services for cooking, lighting, industry, transport. Further, biomass-based energy sources have a tremendous role to play in providing energy services at the village, city or national levels. Hence, the question becomes: is there enough land to emphasize both *energy for biomass production* as well as *biomass production for energy*?

The answer depends upon another question: *Can the biomass-energy feedback loop be closed?* i.e., can the agricultural sector produce sufficient biomass-derived energy to meet its own energy requirements even after adopting the relatively energy-intensive modern biomass output (attributable to the extra biomass-derived energy that is needed for its modern agricultural technologies) sufficient to produce the energy needs of the agricultural sector after supplying the needs of the domestic, transport, etc., sectors?

The biomass-energy feedback loop can be closed only through the establishment of a synergism between the decrease of biomass demand through efficiency improvements and a stepping-up of biomass supply. Biomass for energy necessarily means reduction of biomass-based energy demand (in the domestic, transport and agricultural sectors) through *new* technologies of utilizing biomass energy which do not jeopardize the provision of energy services. It also involves development-oriented variants of modern agricultural technologies for increasing biomass production, and *new* biomass-based energy-supply technologies for the conversion of biomass feedstocks into efficient fuels (producer gas, methanol, ethanol).

Thus, a development-oriented view leads to an approach that is based on a synergism between ENERGY FOR BIOMASS PRODUCTION for a sustainable agriculture and BIOMASS FOR ENERGY PRODUCTION for a sustainable energy system. Only such a synergism can yield SUSTAINABLE RURAL DEVELOPMENT!

Otherwise, a preoccupation with ENERGY FOR BIOMASS PRODUCTION will result in a focus on Food, Fibre and (Non-Fuel) Forest Products with Fuel being forgotten; or an emphasis on BIOMASS FOR ENERGY PRODUCTION will lead to a concentration on Fuel with Food, Fibre and Forest Products being forgotten.

**Seshadri C.V.** (b. 14 April 1930) did Ph.D (1959) Carnegie-Mellon University, USA; He is Director, Shri A M M Murugappa Chettiar Research Centre, Photosynthesis and Energy Division, Madras; and was formerly Dean of Student Affairs, Convener, Council of Deans, IIT Kanpur (1973-74); Assistant Associate Professor and Head of Department of Chemical Engineering, IIT, Kanpur (1965-74).

Seshadri has devoted himself to problems of the rural poor. He has worked on the premise that all rural problems should be tackled *ab initio* and that any solution must be carried out with local participation. He was actively involved in working in the Bihar Famine Areas. Now he is involved in work carried out in Tamil Nadu, Andhra, Kerala, Orissa, W.Bengal, Uttar Pradesh, Maharashtra and Gujarat. His Centre has devised many models and invented instruments, some of them being Anula Windmill, Poghil II Windmill, Jwala Biogas Plants, Chittha Desks, Vatsi Kattumarams, Solar Roof, Solar Regenerator Systems, Wood-ash Soaps, New Energy Quality Markers, Coiling and Helices etc. which have been copied in many countries.

Seshadri is Member of Sigma Xi, USA; Indian Society of Biotechnology; Indian Institution of Chemical Engineers; He is the recipient of NRDC President of India Award for Inventions (1979); World Intellectual Property Organization Award (WIPO) for Invention (1981); Jamnalal Bajaj Award for S & T for Rural Development (to the Research Centre) (1981); Kasturi Ranga Memorial National Award (1983); Professor B.D.Tilak Lecture Award (1986)

**Address:** Director, Shri A M M Murugappa Chettiar Research Centre, Photosynthesis and Energy Division, Tharamani, Madras 600 013; Res. Plot No 75, East Coast Road, Injambakkam, Madras 600 041.

# **EQUITY IS GOOD SCIENCE**

## **PART I. TIME, CALENDAR TIME AND PEASANT CONCEPTS OF TIME**

C V SESHADRI and V BALAJI

*Science and Technology should incorporate the value systems of the local populations in order to ensure that its benefits find a congenial acceptance. This paper, which is in two parts, one on TIME and the other on the ORIGINS OF GENDER, explores the possibility of incorporating native concepts into the axioms of Science and Technology as a means of anticipating equity among peoples as an end product of their applications*

### **INTRODUCTION**

This article begins with a strange hypothesis that may become familiar on acquaintance, that is: the application of science and technology to peasant\* societies may not improve their lot, not only because of often poor implementation, but because these disciplines have axioms and values tied together inextricably that are just not congenial to the society. To most people the idea that science has axioms based on certain values may seem strange. That these axioms and the body of generated knowledge have a non-benign effect on peasant societies may seem even stranger, since the received doctrine is that science is an absolute body of truths based on objective facts and verifiable propositions. Science has rejected concepts of (value-based) absolute space and time, but the idea that science (read: modern science) itself is not absolute but value-based is heresy to the articles of faith of most developing societies. In this article we try to support our hypothesis and hope that at the end of it such support will be self-sustaining. Nowhere, do we believe, is our hypothesis more justified than in the case of TIME.

We first state that our hypothesis rests on two assumptions: (a) it is necessary for peasant development that the people participate in their own development—it is not just a matter of giving devices and money and (b) to

---

The word 'peasant' is used in a generic sense to signify societies that comprise marginal farmers, nomad-fishermen, transhumant communities, hunter-gatherers, tribals and others of their kind. More than 300 million Indians fall in this class.

properly participate, they should comprehend the knowledge system that is being thrust upon them. We believe that the knowledge system will never be comprehensible unless it incorporates a 'native' viewpoint. In other contexts,<sup>1</sup> we have used the description 'equity is good science' to describe the incorporation of native concepts into the axiomatics of knowledge systems. This study is an attempt to do this with TIME.

## ON THE UNDERSTANDING OF TIME

It is a commonplace among practising scientists and indeed most people that time is a one-dimensional continuum that is monotonic (increasing) and has what is loosely termed as an 'arrow'.<sup>2</sup> However, time remains an active field of study and has generated continuing interest.<sup>3</sup> Since Newton and Einstein have provided the basis of our modern insights into time, we shall first discuss their statements before proceeding to our own analysis. It appears to us necessary to point out some misconceptions about their contributions and to clear the way for arriving at a new perspective.

Newton wrote: *"Absolute, true and mathematical time of itself, and from its own nature flows equably without regard to anything external and by another name is called duration; relative, apparent and common time is some sensible and external (whether accurate and equable) measure of duration by the means of motion and this which is commonly used instead of true time; such as an hour, a day, a month, a year"*<sup>4</sup> The keywords here are 'absolute' and flows' and we shall revert to them; in passing we mention that it is the part till the word 'duration' that is most often quoted leaving out the latter half. This we consider unfortunate.

In his 1905 paper, Einstein<sup>5</sup> says [Quotation (1)]: *"So we see that we cannot attach any absolute (emphasis his) signification to the concept of simultaneity, but that two events when viewed from a system of coordinates, are simultaneous can no longer be looked upon as simultaneous events when envisaged from a system which is in motion relatively to that system."* Again, Einstein explaining his own work can be quoted twice.<sup>6,7</sup>

[Quotation (2)]: *"Space and time were thereby divested not of their reality but of their causal absoluteness — i.e. affecting but not affected — which Newton had been compelled to ascribe to them in order to formulate the laws then known."*

[Quotation (3)]: *"By this procedure time lost its absolute character and was adjoined to the "spatial" coordinates as of algebraically (nearly)*

*similar character. The absolute character of time or particularly of simultaneity was destroyed, and the four-dimensional description was introduced as the only adequate one."*

In these quotations, Einstein is explaining, as a sequel to the 1905 paper, what he had done *vis-a-vis* Newton's ideas and we have chosen only a few bearing on time. However for completeness we add one by him on Newtonian space,<sup>8</sup> [Quotation (4)]: *"Therefore in addition to masses and temporally variable distance there must be something else that determines motion. That 'something' he (Newton)\* takes to be relation to 'absolute space' He is aware that space must possess a kind of physical reality .."*

This sample of quotations leads us to the inescapable conclusion that they are using the word 'absolute' in two different senses. Newton is clearly saying that what the clock measures is not time; herein lies the importance of the neglected second part of Newton's words, especially, one should note the reference to 'hours', 'days' etc. There is something else that is 'absolute' time.<sup>9</sup> Measure and duration are used to signify something that is not 'true' time. In other words it seems to us that Newton was concerned about the nature of time and not only its measure.

The only time the word 'absolute' occurs in the 1905 paper of Einstein, Quotation (1) he is referring to the impossibility of defining simultaneity (of clock readings) in 'absolute' terms. And in this first paper on relativity, he has only one concept of time—that which is the same as its measure, namely a clock-face. Unlike Newton he is not concerned with any other quality it may possess. In Quotation (3), it is a bit clearer about what Einstein means referring to 'absolute' time—it is something that cannot be disconnected from space (in the sense of Minkowski)<sup>10</sup> hence leading to the 'only adequate' description. It has no separate and 'absolute' existence apart from the space coordinates. Based on this and several other works of Einstein,<sup>11-12</sup> we think that at no time did Einstein concern himself with time that was not one-dimensional, monotonic, arrow-like. Reading into Einstein a refutation,<sup>13-15</sup> of Newton's idea of 'absolute' time appears to us a misinterpretation. Nowhere in *the Principia* or in his other works,<sup>16</sup> did Newton use time other than as a one-dimensional measure of duration which he acknowledges is not 'true' but this is not the common conception, nor even that of Einstein, as borne out

---

\* '(Newton)' added by the present authors.

by Quotation (2). In other words, Newton has never used his own 'true' or 'absolute' time, but Einstein thinking he had, denies its apartness (of itself and) from space.

Quotation (4) is interesting because we think that Einstein and Newton have converged on the same meaning of the word 'absolute' but only here. The point we are making is not to quibble at semantics that is past, but to strive for sense for the present. Hence to sum up what we have read into them:

(a) Newton recognised that time may not be just what clocks show. Though forced to use clocks, he left the question open.

(b) Einstein was only concerned with time as seen by its measure and direction, that is commonly acknowledged to be proceeding one-way i.e., clock-time or calendar time.

(c) He proceeded to demonstrate that his time did not have independent existence apart from three or more space coordinates.<sup>17-18</sup> In this sense, clock time was not 'absolute' but depended on the motion of coordinate systems.

(d) He did not worry himself about the nature of time

This study is not about the behaviour of clocks, but about the behaviour of time.

## ON THE CONCEPTS OF CLOCKS, CALENDAR-AND ARROW-TIME

Clocks, calendars and arrow-time are mutually reinforcing and since they are fundamental to much of physics and all of engineering, it behooves us to examine their justification in practice. We endeavour to show that these concepts may only be part of the picture that is Time, and that many of the consequences that support and take strength from them may be incompletely understood. As an example, the thermodynamic law of monotonic entropy increase may have to be reexamined if time is not merely arrow- or calendar-like. And all energy calculations and planning depend on thermodynamics.

*"Through the medieval period the cyclic and linear concepts of time were in conflict."* We have just quoted Whitrow<sup>19</sup> who has made a detailed study of the origins of our present concepts of time. Thus Gregory XIII founded the calendar in March 1582<sup>20</sup> and since then, calendar or arrow-time has steadily gained ascendance over cyclic concepts not the least of

whose adherents was Newton.<sup>21</sup> In the rest of this section we shall first delineate the connections between physics and arrow-time; then we shall briefly talk about the interactions of such time and peasant societies before proceeding last of all to examine the geometric connections between arrows and cycles of time. In the next section we advance new constructs of time.

As already stated, time is loosely called an 'arrow'. Davies<sup>22</sup> and Bridgman<sup>23</sup> have questioned this kind of terminology and connect it to a 'psychological one-way movement' or 'an apparently illusory forward flow of psychological time'. Nevertheless, the feeling that 'time is a count-down to death'<sup>24</sup> is pervasive among human beings and undoubtedly forms part of our view of time-asymmetries. A sub-set of this feeling is that intervals of measure-time or elapsed moments are associated, one-to-one, with the number line which goes from  $-\infty$  to  $+\infty$ ; the countability then gives an intrinsic direction or arrow.

Other than such 'non-objective' factors about the anisotropy of time, there have been extensive studies<sup>25-27</sup> about concepts such as irreversibility and the origins of time asymmetry. It is not our intention to cover the literature here but to pose the major questions that are current: (a) the laws of physics behave symmetrically with respect to a one-dimensional time parameter: is the arrow-like behaviour of processes to be found in the initial or boundary conditions or elsewhere? (b) Do the known time-anisotropic processes have a master asymmetry—what are the origins of time? (c) Is clock-time, time?

We discuss the second question first and the first question later. The well-known one-way processes supporting our concepts about one-dimensional time are cosmological expansion, thermodynamic irreversibility, electromagnetic absorption, quantum mechanical interactions and the decay of neutral  $K^0$ -mesons. Whether these lead to our time concept or does time have to be brought in from completely outside these systems is still an open question.<sup>28</sup> A discussion of these problems and suggested solutions is given in Sastry.<sup>29</sup>

We have already stated that time is more than what is shown by clocks. The relationship between reality and measure in physics appears to have led to the following logic: if reality is that which is measured and clocks measure time, then clock-time is real time. The point of departure from such logic is not the 'circle of confusion' that prevails around the edge of every measure, but that 'clocks and calendars are predicated on



'arrown'. Most natural laws are tautological;<sup>30</sup> it can be said that measure is similar in that we design measures for our view of reality, in this case, 'arrows'.

It is also to be noted that: *"The central problem of exact time measurement is to find some periodic cycle that never changes or changes so little that the variations can be disregarded"*<sup>31</sup>-such as the resonance rate of the Cs<sup>33</sup> atom. Time is linear but its measure is cyclical. Is time therefore more than its measure? Can time be arrow-like as well as cyclical? We thus come inexorably to peasant time.

In 1805, Sir Thomas Munro introduced the concept of the Sunday-holiday (1805 Regulation) into India.<sup>32</sup> But peasants have no Sundays. When enmeshed into their cycles of husbandry, they cannot leave it on a weekly basis. Calendar concepts and peasants come into conflict in other ways. Shahid Amin<sup>33</sup> has listed ways in which farm tax collection-schedules based on calendar time have impoverished entire areas of the country. A significant way in which such conflict arises is the school-calendar. The monsoon comes in June when schools start their 'year'. Should the marginal farmer enter his ploughing cycle with his able-bodied children or should he send them to school? It is no surprise that 8 to 10 year old children have the highest drop-out rate in such societies.<sup>34</sup>

What then is peasant-time? We do not know, but we make two statements based on our understanding. Statement one is that it is commonly understood that societies of antiquity and modern peasant societies live on cyclical time. This is basically an attempt to impose mind-body separation and Cartesian categories on people who are essentially a-Cartesian. Concepts of time expressed independent of human existence fall into Cartesian categories; this is a capability that peasants do not have. Statement Two is that the literature of chronobiology<sup>35</sup> shows that most living organisms from single cells<sup>36</sup> to mammals have endogenous 'clocks' that show precise cycles of time where the organism repeats the same sequence of activities in the same order as measured by clocks. We hazard the guess that peasant societies live on some combination of such endogenous rhythms and calendar-time.

Thus far in this study we have made the assumption that one-dimensional time and calendars give mutual support to each other. We now examine the geometric consequences of one-dimensionality when extended to time. We borrow from Mandelbrot<sup>37</sup> the concepts of Euclidean dimension as enunciated by Poincaré<sup>38</sup>: *"When we say that*

*space has the dimension three, what do we mean? If to divide a continuum  $C$  it suffices to consider as cuts a certain number of distinguishable elements, we say this continuum is of dimension one. if on the contrary...to divide a continuum, it suffices to use cuts which form one or several continua of dimension one, we say that  $C$  is a continuum of dimension two...."*

Assume that time is  $n$ -dimensional. Cuts of  $(n-1)$  dimensional 'space' are needed to identify its  $n$ -dimensionality. Assume that time is only one-dimensional. The criterion then specifies that we need zero-dimensional time to identify its geometry. But this violates our hypothesis of time being solely one-dimensional. Also the concept of a distinguishable point is meaningless in one-dimensional time since any present instant instantaneously becomes the past—hence, it becomes indistinguishable.

Using spatial references with time is only part of the problem. It appears that the identification of the dimensionality of time is an unsolvable problem. Consider the following inter-related questions: (a) Can arrow-time be circular? (b) Can arrow-time be double valued? (c) Can there be a reversal of a one-dimensional trajectory, in this case time?

The answer to (a) is that time cannot be one-dimensional and describe a circle,<sup>39</sup> since then we need a time-like point outside the line, the centre. Hence, its one-dimensionality is violated. The answer to (b) is that no one-dimensional 'line' can approach arbitrarily close to any point on itself without enclosing an 'area' and (c) requires a plane of reflection or an axis of rotation which are themselves time-like to allow us to substitute  $(-t)$  for  $(t)$ , (time). These then violate our assumption of one-dimensionality.

The answers to (a) and (b) show that the perception of an 'eternal return'<sup>40</sup> as being equal to the cyclical behaviour of linear time<sup>41,42</sup> is the result of a non-sequitur: indeed, it is manifestly absurd. In simpler terms, cycles of peasant-time cannot have reference to clocks.

Clocks and time-machines that take us back to some bygone era are mutually exclusive. The equations of motion, according to (c), are not symmetric in time that is one-dimensional—dimensionality and time-reversal are not compatible concepts.

Gal-or states<sup>43</sup> that it is logically absurd to convert cyclical time to linear time. We have shown that the reverse is also not possible. Gödel in

a well-known paper<sup>44</sup> solves the cosmological equations to show that the trajectory winds back on itself—i.e. a return to the past takes place. This as we have stated already is something that cannot follow from the premise. We have not been able to trace any reference in the relativistic literature that discusses anything other than one time-dimension.

### ON A NEW CONCEPT OF TIME

We have shown that the word "flows" in Newton's original quotation should not be thought of as referring to something similar to the flow of a fluid. We also stated that we cannot describe peasant-time as cyclised one-dimensional time; instead, it is best to view it as a combination of cycles (rhythms) and calendar-time. Our construct of TIME starts with this realisation.

We define Umbon-time\* as being some linear measure of time. We define Kaalon-time\*\* as cycles of time. Our model combines umbons and kaalons such that umbons are to be considered as a measure that is used to relate to present clock-time. The overall character and behaviour of time should be described in terms of kaalons. This is so because all the events around us occurring in Nature are cyclical and even the measurement of clock-time demands the existence of invariant cyclical events.

We describe one possible combination of kaalons and umbons in Fig. 1 with the understanding that it is a 'spatial' representation. It is constructed as follows: on the largest circle with centre at 'O', six circles are drawn such that (i) their centres are on its circumference and (ii) any one of them touches the neighbouring two. On each of these circles, the same process is repeated, till the smallest circle is of infinitesimal radius. Four iterations drawn on two such circles are shown in this figure. For the sake of clarity only part of the figure is shown completely. The construction can be extended in all directions from 'O' (See appendix for this and other arrangements).

The figure shows that kaalons (cycles represented here as circles† behave similarly with respect to each other and with respect to a reference

---

\* Umbon: from the Tamil word *Umbu* for arrow.

\*\* Kaalon: from the Sanskrit word *Kaala* for time.

† Any closed curve would serve as a representation for kaalons but circles are chosen here for ease of construction.

point at all levels. This figure also has some resemblance to fractal curves,<sup>37</sup> though, as we have stated, applying concepts of dimensionality to time is fraught with uncertainty.

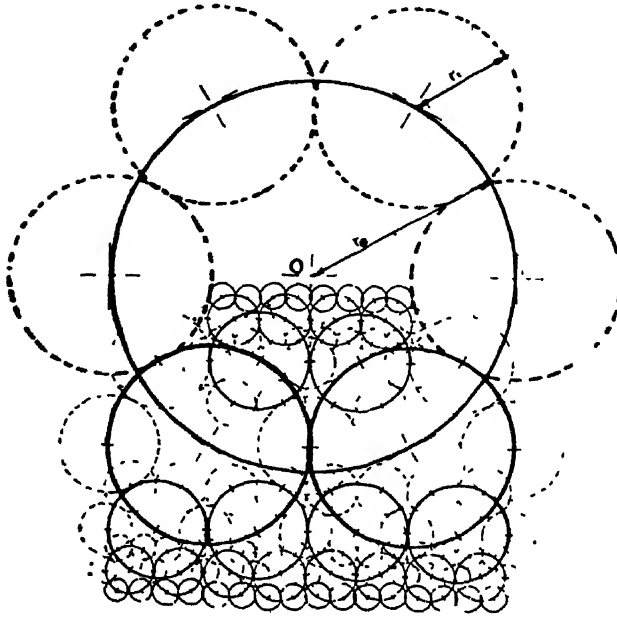


FIG 1 Two-dimensional spatial representation of kaalon-time

We now use the feature of self-affinity of the representation to describe umbon-time (Fig. 2). Three points P, A and B are shown in this figure. By the very nature of representation, all these points will fall on arcs of circles (representing kaalons). Point A can be reached from point P by moving along various arcs in any direction (ultimately, this would involve covering a very large number of circles, because of the self-affine nature of the construction). Once A is reached, a straight line PA gives the 'distance' covered, averaging out the total distance along the arcs. This measure is umbon-time. Similarly, PB measures umbon-time. The distance AB does not have intrinsic relevance, unless viewed as movement from PA then B or PB then A; in that sense, difference of PA and PB may be thought of as providing an umbon-measure of 'time-lapse' between events A and B. But such a measurement necessarily requires reference to a point on a cycle, such as 'P'.

We now have to justify the existence of kaalons as the fundamental nature of time. Short of inventing a kaalon-chronometer or determining a

fundamental kaalon-unit, kaalon-time cannot be placed on a quantitative basis. We can only conjecture further why time is so.

The endogenous rhythms of living things can be considered as repetitive cycles of kaalons, without reference to external clocks. Indeed all living things repeat the origins of life—if life is assumed to begin with the association of DNA, RNA and the proteins of the replicating units.<sup>45</sup> In doing so, they have to go through many millions of cyclical steps e.g., DNA synthesis, replication, transcription, translation etc.,<sup>46,47</sup> which is impossible to imagine merely as a recall of events from a storage type of memory-device because such storage would be just too large. Since the representation (Fig. 1) is self-affine, it is possible to find a point on a kaalon-cycle, which is congruent to or identical with any chosen reference point (the origins of life). Therefore, recall of origins during the process of DNA synthesis, replication etc, does not involve reaching the actual origin backwards in clock-time but to reach an equivalent point. Thus, such a process of recall avoids the need for an impossibly large memory device. By its very nature, kaalon-time is also memory. The averaging-out of all the kaalon-trajectories gives the umbon time for the process.

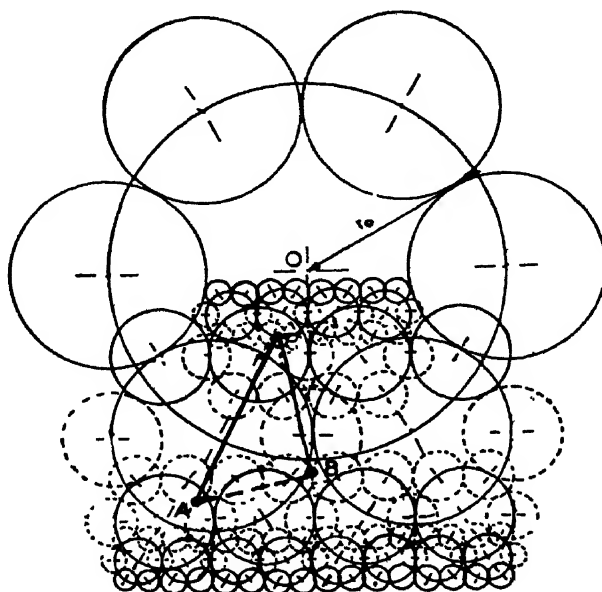


FIG 2 Kaalon-time and its umbon-measure

The consequences to areas such as physics and chemistry if time is really kaalon-like need to be explored further. Here we have used an

example from biology to show why time is not merely umbon-like. In the words of Synge<sup>48</sup> "... *however much they may have been inspired by nature, mathematical theories are no more than maps or models of nature.*" In the case of time, Science seems to have restricted itself to a model of Nature that unlike waves, particles and energy is singular in character and uniquely 'one-dimensional'; this stems from the beliefs of its practitioners that are not necessarily global. Here we have speculated about its having a dual nature, namely kaalon-like and umbon-like; the mathematical modelling is yet to follow. Such a view of nature may be less known but is more prevalent.

### CONCLUSION

We began this article pointing out that scientific concepts, such as that of time may have to be re-examined in applications to peasant societies. We tried to show that concepts such as the 'arrow-like' behaviour of time and its dimensionality continue to be nebulous and that new constructs may be necessary. One such construct combines cycles or rhythms called kaalons and arrows called umbons in a new fashion; though crude and as yet untested, such conjectures may help us understand the intrinsic nature of time better.

### ACKNOWLEDGEMENTS

One of us (CVS) is grateful to the Indian National Science Academy for an award (B D Tilak Award, 1986) as stimulus for this paper. Parts of this and a companion study (part 2) were presented at the INSA Award Lecture. He is also grateful to the Kasturi Ranga Memorial National Award Committee (1983) for financial aid. VB is the recipient of a CSIR Senior Research Fellowship. S Sastry and J Thomas have provided valuable insights and documentation.

### References

1. C V Seshadri 'Equity is good science' Paper delivered on the occasion of the 125th Anniversary celebration of the University of Madras, September 1983; 'Equity is Good Science' Paper presented at the *Int Symp Hunger, Poverty and Disarmament Moscow*, (1986)
2. A S Eddington *The Nature of the Physical World* Everyman's Library London (1935) 75

3. J T Fraser *et al* (Eds) *The Study of Time* Springer-Verlag New York and Heidelberg (1972-81) 1-4
4. Isaac Newton *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World* Cambridge University Press Cambridge (1934) 6
5. A Einstein 'On the electrodynamics of moving bodies' *Ann Phys* 17 (1905); In H A Lorentz *et al.* A Einstein H Minkowski and H Weyl *The Principle of Relativity* Dover New York (1952) 42
6. A Einstein 'The mechanics of Newton and their influence on the development of theoretical physics' *D Naturwiss* 15 (1927) reprinted in A Einstein *Ideas and Opinions* Rupa New Delhi (1981) 308
7. A Einstein Physics and Reality *J Franklin Inst* 221 (3) (1936) reprinted in A Einstein *Ideas and Opinions* Rupa New Delhi (1981) 260
8. AS Einstein *Ibid* 258
9. I A Akchurin and M D Akhundov 'Einstein and the development of the concept of space' In: *Einstein and the Philosophical Problems of 20th Century Physics* Progress Publishers Moscow (1983) 147
10. H Minkowski Space and time in H A Lorentz *et al* *The Principle of Relativity* Dover New York (1952) 80
11. A Einstein *The Meaning of Relativity* (5th Edn) Princeton University Press Princeton (1955) 29
12. A Einstein and L Infeld *The Evolution of Physics* Cambridge University Press New York (1961) 208
13. G J Whitrow *The Nature of Time* Penguin Middlesex UK (1973) 95
14. P C W Davies *Space and Time in the Modern Universe* Cambridge University Press (1974) 31
15. P G Bergmann 'The quest for unity: general relativity and unitary field theories' In: *To Fulfill a Vision* Y Ne' eman (Ed) Addison Wesley Reading (1981) 13
16. Isaac Newton *Optics* published as Vol 34 of *Great Books of the Western World* William Benton Chicago London Toronto and Geneva (1952)
17. A Einstein *The Meaning of Relativity* (5th Ed) Princeton University Press (1955)
18. A Einstein *Ann Phys* 49 (1916); reprinted in Lorentz *et al* *The Principle of Relativity*, pp. 111-120
19. G J Whitrow *The Nature of Time* Penguin Middlesex (1973) 19
20. G J Whitrow *ibid* 11
21. G J Whitrow *ibid* 22
22. P C W Davies *Space and Time in the Modern Universe* Cambridge University Press Cambridge (1974) 59

23. P W Bridgman *Reflections of a Physicist* Philosophical Library New York (1950) 163
24. C V Seshadri 'Equity is Good Science' Paper presented at the International Meeting on Hunger, Poverty and Disarmament Moscow (1986)
25. T Gold and D Schumacher (Eds) *The Nature of Time* Cornell University Press Ithaca (1967)
26. B Gal-or *Science* 176 (1972) 11
27. P T Landsberg et al (Eds) *Pure and Applied Chemistry* 22 (1970) 412 et seq
28. J A Wheeler *American Scientist* 74 (1986) 366 et seq
29. S Sastry *The Concept of Time in Physics, a Review Report* submitted to CSIR New Delhi (1986) MCRC Madras
30. C A Truesdell and R Tuopin The classical field theories In *Handbüch der Physik* S Flüge (Ed) Band III/1 Springer-Verlag Berlin (1960) Eq. 157.1
31. L Coe *Am J Phys* 37 (1969) 810
32. V S Sastry *Thomas Munro* Madras (1924) XXVII
33. S Amin Small peasant commodity production and rural indebtedness: the culture of sugarcane in eastern UP, c 1880-1920 In: *Subaltern Studies* I R Guha (ed) Oxford University Press New Delhi (1982) 84-85
34. K Ramakrishnan *Adult Education in Tamil Nadu* Madras Institute of Development Studies, Seminar Papers (1985) Madras
35. *Cold Spring Harbor Symposia on Quantitative Biology "Biological Clocks"* XX The Biological Laboratory Cold Spring Harbor Massachusetts (1960)
36. C H Johnson and J W Hastings *Am Scient* 74 (1986) 29
37. B B Mandelbrot *Fractals: Form, Chance and Dimension* W H Freeman San Francisco (1977) *The Fractal Geometry of Nature* W H Freeman San Francisco (1984)
38. H Poincaré *Science and Hypothesis* Dover New York (1952) 32
39. B B Mandelbrot *Fractal Geometry of Nature* W H Freeman San Francisco (1984) 446
40. F J Tipler (Ed) *Essays in General Relativity: a Festschrift for Abraham Taub* Academic Press New York (1980) 21
41. G J Whitrow *The Nature of Time* Penguin Middlesex (1973) 99
42. P C W Davies *Space and Time in the Modern Universe* Cambridge University Press cambridge (1974) 198
43. B Gal-or *Cosmology, Physics and Philosophy* Springer New York (1981) 448
44. K Gödel *Rev mod Phys* 21 (3) (1949) 447



45. C V Seshadri and N Jeejibai Conjectures on the dual origins of the male and female *Proc INSA Part B, Biological Sciences (to be published)*
46. A Kornberg *DNA synthesis* W H Freeman San Francisco (1974)
47. J D Watson *Molecular Biology of the Gene* Benjamin New York (1975)
48. J M Synge Classical Dynamics in *Handbüch der Physik* S Flugge (Ed) Band III/1 Springer-Verlag Berlin (1960) 3

## APPENDIX

Fig. 1 discussed the case of six circles (of kaalons) with centres on the circumference of each other circle. The construction is quite straightforward and it can be shown that the sum of all the radii of each-sized circle i.e.,

$$r_0 + r_1 + r_2 + ..... + r_n + ... = 2r_0.$$

This is because each succeeding radius is  $1/2$  the preceding one. If the construction had started with four circles instead of six, the sum of the individual radii can be shown to be  $(2+ 2) r_0$ ; in fact, any number of kaalon-circles greater than two on the first circle can be shown to have a sum of radii that converges. Other construction that do not have the centres of the smaller circles on the circumference of the larger circle can be made and shown to have similar convergences. A large number of constructions of intersecting circles can be shown to be similar to fractal curves. The significance of the convergence limit, if it exists for kaalons, has to be explored.



**Shrinath Sheshgiri Kalbag** (b. 23 October 1928) did Ph.D. (1955) from University of Illinois, USA; He is Director, Centre for Science Education and Research, Pabal, Pune (1982- ); and was formerly Manager of Technical Services, Hindustan Lever Ltd. (1963-82).

Kalbag has produced several patents including for perfumery chemicals, liquid-liquid extraction of oils by continuous centrifuges, ester sulphonate detergents, microporous Ni catalysts, and microprocessor controlled Industrial Scale Chromatography on 5000 tons/annum scale which have

been transferred for production on a commercial scale. Later Kalbag took up a project, supported by DST on Non formal Science Education in Pabal, near Pune. He has developed a system of science education, that integrates education with rural development based on 1) Learning by doing 2) Problem solving orientation 3) earning while learning 4) Community paying for services received. This system is now being implemented in one school in a formal way and also given to school drop outs in a non formal way.

Kalbag is the recipient of Prof. B.D. Tilak Lecture Award (INSA) (1987).

**Address:** Director, Vigyan Aashram, Centre for Science Education and Research, Pabal 412 403; District Pune; Tel: 26.

## **SCIENCE & TECHNOLOGY FOR RURAL DEVELOPMENT-IS IT DIFFERENT ?**

**S S KALBAG**

We are all aware of the substantial investment that we as a nation have made in S & T education and research. It will also be generally agreed that the benefit from this huge and continuing investment do not flow to all sections of our society in equal or fair measure. While part of our nation lives in the late 20th century, another part is still living in the medieval ages. No nation can progress much beyond the "national average". When these differences are too wide, they create not only social tensions but slow down progress and make it expensive, as we are now finding.

How can the whole nation benefit from whatever S & T we have invested in ? This is the question that worries not only the politicians but also the scientists. There is one group that feels we must set up separate establishments for "Appropriate" technology for Science for Rural Development, etc. There have been demands for establishment for "Vedic Science" "Indian Science" or "Indian Medicine", etc. While there is a case for funding organisations for specific objectives, I think we should not consider "Appropriate Technology" or "Ayurvedic Medicine" as a separate kind of Science with different basics, where normal scientific principles are not valid.

Terms like "Appropriate Technology", "Renewable Energy", etc. in relation to use of S & T for Rural Development have produced some distortion in our efforts. It is my objective to try and show how the methods of S & T investigation used in basic and industrial research are badly needed for the development of the Rural India.

How can we get more S & T effort channelised to developing the backward regions and sections of society ? Is it only necessary to allocate more funds specifically for that ?

Briefly my thoughts in this regard are:

- (i) We should spread scientific methodology through education to all sections of society, in a way so that it is absorbed into the culture.

- (ii) We must have a system which poses the right questions/problems to the right people.

If these things can be done, I believe a lot of benefits will flow even from the existing S & T establishments to rural societies. What is more, this will elevate the overall quality of our S & T effort and make it more original and less of a blind follower of the West.

There are two ways in which we apply S & T for progress. First, we have some novel technique or product that fascinates us. We are, therefore, on the look out for an opportunity to use this information to our greatest advantage. Here, solution of the problem is secondary and a successful application of our "idea" is the main objective. Most of our effort in S & T application falls in this category. In the second situation, we have a problem that seriously bothers us. We look at it, analyse, diagnose, think of alternatives, and select the best idea to meet the required situation. Relatively less effort goes into this kind of problem solving.

Both methods are valid, proven tools are complementary. The overdependence on one, however, will impede progress. My belief is that at present we mostly look for "problems" to fit known solutions, but not the other way round. Primarily this is because the "problem" awareness in the scientific community is biased by poor information flow from our rural segment. Therefore published scientific information from (mostly Western) Journals becomes the major inspiration for selection of new projects for investigation.

We have to select problems to tackle on some basis—because we have finite, limited resources. This is therefore a crucial decision. What are the scientists' expectations? I believe every scientist hopes for recognition and appreciation by his peers. Recognition comes from success in a challenging situation. If the work is trivial, success is not valued. If it is too difficult, success is elusive. A breakthrough is what makes the scientific work respectable. I am suggesting the following guidelines for selecting projects which I believe will meet the aspirations of the scientist and the needs of the society.

- (i) We should be able to utilise our strength – viz technical skill. There must be a challenge. The exercise must sharpen our skill and we must thus "grow" by solving or even attempting to solve that problem.

- (ii) If solved, the program should experience exponential growth. All successful products of technologies show a sigma type growth curve. If our solution depends on charity or subsidy, the growth will be at best linear.
- (iii) We should look at on going activities for "problems" to tackle. They usually have relevance and urgency. Even small improvements have big impact.
- (iv) We should involve costing and economics from selection stage of the project. Without cost consciousness development efforts can go astray.
- (v) All activities have bottlenecks—limiting parameters. When these are tackled there is progress and new parameters become limiting. Removing bottlenecks is therefore a continuing activity that ensures continuous progress until the law of diminishing returns make it unattractive.
- (vi) Drudgery is an evidence of the crucial nature of the activity. Drudgery is removed by improving productivity. This also reduces the cost. Measurement of productivity of every operation therefore brings out areas needing development.
- (vii) Any project must have a well defined "user Population". The specification of this population defines the boundaries within which we have to find our solution. Our sights should always be fixed on this population for all decision making.
- (viii) We often get confused by the issue of Capital vs Labour. The only true Capital is the Human Resource and time. Money is only a token of assets built-up in the past through human resource and time + natural resources. Our choice of technology should be based on what combination of available resources produces the fastest growth of assets.
- (ix) We should not be biased by luxury vs necessity arguments. Electricity was a luxury 100 years ago, now it is a necessity. If something has the potential to make an impact on the quality of life, S & T can take up the challenge to bring it within the range of every one's purse, and make it a necessity.

- (x) Finally, we should select a project that has a reasonable chance of success. This is necessary for morale as also for keeping open the flow of funds.

The criterion for success of the project lies in its success in implementation—the success in the market place; not on laboratory test results. S & T is a hard task master. Unless all the "nuts and bolts" are properly tightened it does not yield results.

The information system is the nerve centre of any S & T programme. It is not often realised that the information system is as important for the selection of the problem as for solving it. The foundation of any information system is the source and quality of information. If we are going to solve the problems of the rural society then this section of the society will be the ultimate source of our information. It is a fact that at present this source is not capable of giving relevant information in usable form. Such information now comes from estimates and is usually old, obsolete, incomplete and often unusable. The first requisite for our information system is, therefore, a population educated in the methodology of science, viz. trained to observe, measure, record, classify, compare, and interpret simple information. This in my opinion is the true scientific temper. If this is absorbed by even a section of the youth in the villages, we could build on that reliable information system that will define, quantify and document all "problems" perceived at the village level. A referral system is then envisaged, whereby these defined problems go to an appropriate S & T establishment, where the level of knowledge and the skill match the challenge posed by the problem. Backing up this referral system should be a computerised information storage and retrieval system, that collects and updates on a continuing basis, technical and economic information, performance indices and make this available to anyone on demand.

Such a system has been elaborated and is under consideration of the Govt. of India. If and when accepted, it will be possible to bring before the scientific community a range of problems, along with reliable background data, so that they can pick up challenging tasks of relevance to the rural society.

I have no doubt that when the awareness of opportunities for the application of good science spreads among the scientific community, the complexion of the scientific research in our country will change to bring greater benefits to our own rural societies. Let us remember that Louis

Pasteur laid the foundation of microbiology by helping to solve the problem of rural France — wine making. Similarly investigation of the causes of malaria of India lead Ross to world class research. Embryo transfer in cattle, tissue culture, new plant strains through genetic manipulation, new techniques for ground water exploration and development, meteorology and a host of other fields hold promise of significant impact on our rural economics.

A key component of the above scheme is to educate and develop the human resource — the rural youth and to bring them into the mainstream of Science and Technology. For this, we have tried the following strategy: (i) work in live situations — clear targets and continuous feed back; (ii) train tutors in villages, who then teach others — second order tutors; (iii) take simple problems first and "grow" with experience. For this our model is the following: (a) Education to be the cutting edge of Development Learning by doing, (b) Multiskill training, (c) Problems solving orientation, (d) Education integrated with community service.

Towards this end, we have evolved a "Rural Technology" course programme that gives a broad spectrum technical education and gives access to the world of modern technologies, to the rural youth. This course is now recognised by the Board of Secondary Education, Maharashtra and is given as a one-year course to school drop-outs and a three-year course for VIII, IX and Xth Standards in the formal schools.

Site: Pabal, is a village of about 3000 in a drought prone area of Pune District. Not too near, not too far, about 53 km to the north, but 20 km away from the nearest highway.

- (i) Water-map reading, understanding geology is basic. Geophysical testing can be demystified and village youth can be used to give a useful service to the farmers — water prospecting.
- (ii) Workshop skills are very strategic in the modern age. Welding, drilling, grinding, turning are not only needed for repairs and maintenance but also to develop the inventive ability and uncover talent in the youth.
- (iii) *Energy and Transport*: Diesel engine maintenance, fabrication of pneumatic wheeled carts and design of a diesel cart as a farm vehicle for transport, pumping and agricultural operations are some of the projects undertaken.

- (iv) Welding forms such a key operation that trainees have brought their own material, built a welding transformer, for less than Rs. 2500/- and taken it back to their village to start their own workshop.
- (v) Engineering drawing is an important component of training. It stimulates visualisation and is the universal language of technology.
- (vi) Low cost housing and sanitation is another subject of training. All the campus structures and experimental and erected through trainee and staff effort. Steel frame houses with ferrocement panels is one design in low cost housing. Geodesic dome structures covered with ferrocement is another alternative. The steel kit for this is available for Rs. 2845/- for the 280 sq. ft. dome (completed cost Rs. 10,000/-). Low cost sanitation forms part of the construction practicals.
- (vii) Agriculture includes pest control, drip irrigation and use of hybrid seeds.
- (viii) *Animal Husbandry* : Poultry is the main theme here, besides cross breed goats and cows. TRYSEM courses are also held. In the poultry training chicks are reared in deep litter and the cage system and the profit is shared by the trainees. Special modular 100 and 200 bird poultry houses have been designed and tested for small poultries. Artificial Insemination in goats is practiced (collection of semen + insemination)
- (ix) Knitting and sewing attract a number of girls. They also get health and child care education, as part of the Rural Technology course.
- (x) All staff live in the campus and eat together. This is part of their education. The staff have a weekly seminar on technical subjects and have a review meeting for work planning and budgeting.



**Pramod Karan Sethi** (b. 28 November 1927) did MBBS, MS; and was awarded FRCS, UK. He is Emeritus Professor of Orthopaedics. S.M.S Medical College, Japiur; Hony Consultant Orthopaedic Surgeon, S.D.M. Hospital, Jaipur; and was formerly Professor of Orthopaedic and Director, Rehabilitation Research Centre, S.M.S. Medical College, Jaipur.

Sethi's earlier work dealt largely with some common clinical problems encountered in orthopaedic practice. He worked out a simple technique of bone grafting for delayed union or non-union of fractures which was effective and practical within the limited facilities available in average hospitals. Later he designed artificial lower limbs and appliances for poliomyelitis specially suited to Indian Conditions. He devised a new hybrid design of a socket, fashioned in an aluminium limb which provided all the advantages of a modern PTB socket and yet permitted adequate ventilation to prevent the skin of the stump from becoming unbearably hot.

Sethi is member (Past Vice-President) Indian Orthopaedic Association; Prosthetics & Orthotics Society of India (Past President) He is the recipient of Magsaysay Award (Manila) (1981), Giants International Award (1985); Dr. L.H. Lobo Memorial Oration (Ludhiana) (1985), Instrument Society of India (National Award) (1986); Dr Ernst Borges Memorial Oration (Tata Memorial Hospital); N.D. Diwan Memorial Award, NASEOH, Bombay (1987); Dr. K.N. Lal Memorial Oration Award (Indian Medical Association) (1987); Dr. P.N. Wahi Oration Award, Delhi (1988), Padma Shri (1981); Prof B D. Tilak Lecture Award (INSA) (1988).

**Address:** Vivekanand Marg, Jaipur 302 001; Tel: 74682

## **DESIGNING AIDS FOR PHYSICALLY HANDICAPPED IN DEVELOPING COUNTRIES**

**P K SETHI**

Since 1981, which the United Nations declared to be the "Year of the Handicapped," there has been, at long last, a growing awareness about the disabled population in our country. Special census have been carried out on the disabled, and their staggering number in our vast country, with 80% of its people living in inaccessible rural areas, has posed major challenges to our planning bodies. But we are passing through some heady days in our country and the numerous problems which beset our society are believed to be solvable by mere technological and managerial interventions. Technology Missions are being set up to speed our entry into the 21st Century because firm promises have been made. This has set the pace for some hectic activity since targets have to be achieved. 'Targeting' is the new word. We need more targeted research, more mission-oriented science. This is said to be the new drift. There is little time to waste in rediscovering the wheel and so it is considered prudent to buy technology packages from the west. This is affirmed as the speediest path to tackle the problem of the vast numbers of disabled in our country.

I sense serious trouble when we initiate such a "top-down" move. Such moves require a much greater store of usable information, with coherence and connectedness, than actually exists. We presuppose that we know what the needs of the disabled are, and all that is required is to activate our already existing centralised production agency for rehabilitation aids, prepare managerial flow-charts to set up an efficient distribution system which can reach the remotest areas in our vast country and use the existing government machinery, with some NGO's thrown in, to assist the implementation of our schemes and another problem would be solved. Statistics are being reeled out to demonstrate the success of such a move.

There is one factor which does not figure in all this hectic activity and that is the disabled persons themselves as human beings. We have never really bothered to find out what their felt needs are. Are they really using the appliances we are handing out to them. What is the incidence of a drop-out rate? If the appliance is not being used, is it because our people

are ignorant and do not realize what is good for them, as is being affirmed by many specialists, or is there a possibility that our appliances are not suitably designed for them? Can there really be a standard universal solution for a particular locomotor deficiency or dys-function or do we need a more culture specific and location-specific alternative? Is it fair to offer them only one design option, take it or leave it, or is it better to work out a range of options from which a particular individual has the possibility to choose what suits him best? Doling out such aids as if the disabled are objects of pity or charity or a mere statistic is a demeaning business and instead of making the beneficiaries more self-reliant, which ought to be the purpose of such aids, one can often erode their pride and confidence.

Raising such questions is always awkward and one is likely to be misconstrued as an obstructionist. It is my submission that the central figure in all such activity has to be the user himself. If the appliance actually helps him or her, it would certainly be used. If not, it will be rejected and good luck to him/her. After all it is the user who knows best what is good for him/her.

Designing such aids, I have learnt, is a very complex business. It not only calls for more science but a much better understanding of our society, its culture, its gross economic disparities and its stratified structure. One would then realize that there is a lot more of basic work which needs to be done and even though this would necessarily be time-consuming, it is likely, in the long haul to be more cost-effective and more appropriate for our target group. Such work, of course, must obviously take into account the pressing needs of our deprived and marginalised people, and *reaching the largest number in the shortest time has got to be one of the objectives of such research.*

I would like to illustrate some of the complexities of this problem by using two examples. One involves an appropriate design of artificial lower limbs for our amputees and the other deals with the polio problem which still remains the largest single cause of physical disabilities in our country.

It is almost exactly three decades from now when I got involved with the problem of providing physical aids to many of my patients. I could not amputate a limb and then wash my hands off by directing the patient to fend for himself in trying to get an artificial limb. Likewise, there seemed to be little point in correcting neglected deformities in children affected by poliomyelitis, performing multiple operations to

straighten out the lower limbs and then advising them to get braces which were then required to support their otherwise flail limbs. From where could they secure these appliances? At that point of time, the only available facilities for such appliances were located at Bombay or Poona. While the affluent could go there, the majority of my patients were poor and the advice to travel a thousand miles away for securing an appliance was totally unrealistic. I realized that without a neighbourhood facility, the bulk of the disabled would remain deprived of aids they desperately need.

This impelled me to organize a workshop in our hospital where such appliances could be made. I somehow succeeded in achieving this objective and I became rather proud when we started making artificial limbs and braces locally.

Having been a product of western education and with a psychology which was heavily influenced by a colonial heritage, my ambition was to see that our appliances were as good as anywhere in the west. Of course, all I could achieve were "blurred xerox copies" of the limbs made in London or New York. In spite of this, I seemed to be satisfied with the progress made.

My initial elation soon received a setback when I started encountering some of my amputees reverting to their crutches. Whenever I encountered such a situation, I questioned them. "Why are you not using the limb we made for you", I would ask. My suspicion was that there was some technical flaw in them; the socket might be hurting their stump or else something must have broken down. But the feedback I got was something completely unexpected. It became obvious to me that I was taking a very simplistic, almost naive approach to the problem of limb substitution. I was taught in the medical school that the function of the lower limbs is to be able to stand and walk on them. I now realized that there are many other attributes in our lower limbs which are equally important to our people but which the western-designed limb did not cater for.

## **TWO CULTURES – FLOOR-SITTING VS CHAIR-SITTING**

In the cold climate of Europe or North America, the feet have to be protected from cold by using warm socks and closed shoes. One has to move away from the cold floor and design a chair to sit on. A table then becomes their work-surface. Also, in most advanced countries, people walk on paved floor and level streets and the foot is not required to adapt

to uneven surfaces. On the other hand, our warm climate makes a closed shoe uncomfortable and most of our people walk barefoot or else in open, well-ventilated footwear, often on the rugged terrain of our countryside where the suppleness of the foot becomes a vital attribute to adapt itself to uneven surfaces. We use the floor for squatting, sitting cross-legged, working, eating or sleeping on it. And so there is the social custom of removing one's shoes on entering a home, or a kitchen, or a place of worship. This is a sensible and hygienic thing to do to prevent dirtying the floor.

It is important to distinguish thus between a chair sitting and a floor-sitting culture because, as you would see, there are important design implications involved.

### THE WESTERN DESIGNED LIMB

One of the important features which characterises a western limb is its footpiece. It is not shaped like a human foot. Instead, its shape is such that it can easily slip into a shoe, which then hides its odd appearance and also protects it from damage. A closed shoe, in other words, is an integral part of the limb design. Take the shoe off and you cannot use the limb.

Providing this kind of limb to our amputees, therefore, made it compulsory for them to wear shoes to be able to use it. You can easily appreciate how such a simple demand can lead to major problems when closed shoes are not only uncomfortable in our hot climate but because they have to be repeatedly removed in a floor-sitting culture. Our women would not agree to wear such shoes anyway and in a rural environment, one cannot expect a farmer to wade through water and mud wearing a pair of expensive Oxford shoes!

Not only this, squatting requires a range of mobility in the knee and ankle which is not available in a western limb. Likewise, the foot is twisted inwards when sitting cross-legged on the floor. The western footpiece, which is otherwise a very clever design, has a solid wooden keel, which prevents any movements at the ankle. So the patient cannot squat on the floor. An attempt to sit cross-legged presses the stiff footpiece, which in turn forces the entire limb to rotate, causing unbearable pressure on the stump. The upshot is that an amputee using a western limb has to remove it repeatedly several times a day when entering his home. When he works sitting on floor, he takes the limb off and then has to use crutches to be able to move around. And so, unless the

amputee changes his life style into a shoewearing, chair-sitting culture, this artificial limb disables rather than helps him. My colleagues keep on telling me that our people are uneducated, stupid and stubborn. The fact is, and I have repeatedly learnt this lesson, that our people are not irrational. They are perfectly capable of making rational decisions. It is we, because of the blinkers we wear, and our lack of sensitivity, who are unable to understand the rejection of our solutions.

### EVOLUTION OF JAIPUR FOOT

So the first item we took up was to redesign the footpiece. A set of desirable functions was listed out. The footpiece should not require a shoe to hide it and protect it. So it should look like a normal foot and be made of a material which is not only flexible but also tough, abrasion-resistant and waterproof. The internal design should provide adequate mobility to enable sitting on floor and walk on uneven ground where the foot is required to adapt to the rugged terrain of our countryside and yet the foot should offer a stable support while walking.

We decide to use a solid rubber elastomer as the outer casing for our footpiece. Several reasons prompted us to choose this material. Solid rubber has many unique properties combining flexibility, toughness, and abrasion and tear resistance. A material which is durable enough for an automobile tyre should be adequate for our footpiece. The material is readily available in our country and an extensive trade in retreading tyres has made our people familiar with vulcanization.

To reproduce the shape of the foot, a 4-piece aluminium die was prepared locally by our traditional craftsmen, who used the age-old sand-casting methods and the cost of this mould was a fraction of what the fancy die-designing firms were asking. By packing rubber into the die and vulcanizing it in our hospital autoclave, a footpiece resembling a natural foot can be obtained.

But mere appearance is not enough. The desirable range of mobility must be available to provide the activities already listed out. Our first sample was made of solid rubber and it was so heavy and stiff as to be totally unusable.

To reduce the weight of this foot, it occurred to us to place a western footpiece into our aluminium mould and then fill the remaining space around it with solid rubber. This encapsulation substituted for a built-in-

shoe which resembles a natural foot. The footpiece was now much lighter and became suitable for barefoot walking.

However, the problem of mobility still remained. The main obstacle was the wooden keel of the western footpiece which prevented squatting on floor. We tried to tinker with the keel, cutting wedges into it to provide mobility but these proved inadequate. Our minds were still wedded to the conventional design and it is so easy matter, you would agree, to move away from our preconceived ideas.

Repeated failures ultimately forced us to make a fresh start and finally we arrived at a completely new design concept. For the ankle region a block of wood had to be provided for securing a carriage bolt which connects the footpiece to the leg. The front part of the foot(foot) also had a separate block of wood to provide stiffness when this part of the foot is loaded when the heel is off the ground. Between these two rigid wooden blocks, a large microcellular rubber block was interposed and this behaved like a universal joint, with a freedom of movement in all directions.

Now squatting was possible.

We had tested our footpieces in the engineering college laboratories and characterised its behaviour under different loading conditions. Field trials on amputees revealed that we could meet all our design criteria. The foot was shown to be very strong, breaking up only under a vertical loading strain of two tons. We were happy.

What we did not take into account was the problem of fatigue as well as disaster failure. Soon amputees started returning with the external shell of rubber cracking open and the internal components virtually spilling out. Then it was suggested that we use reinforcement with rubberized tyres cord which is used in car tyres to prevent such disaster failures. We had then to become familiar with this new material, learn how to lay out the re-inforcement so that the desirable range of mobility was not adversely affected and we ended up with a product which had a durability span of 3 to 5 years under tough field trials in rural areas.

Periodically, amputees would come back to us with cracks in the footpiece. The curious thing we noticed was the almost consistent location of distribution of these cracks. So we started plotting the cracks in the damaged footpieces. These were always around the ankle region. The moment we realized this, the reason become obvious. The encapsulation

around the wooden blocks was immobile. The entire mobility resided in the junctional area which was in the hindfoot region and this was the place where all stress concentration was located.

A major design revision was then made and we replaced the forefoot wooden block with another MCR block of a higher Shore hardness and appropriately stiffened with the tough tyre cord. This has resulted in a more uniform distribution of stresses spread over the entire hindfoot and forefoot. Not only has it added to the life of the footpiece but it has provided us with an additional bonus of the forefoot gaining an independent range of mobility. This allows for a much better adaptation of the footpiece to uneven surfaces.

We realized then that what started off with a relatively culture-specific design need paid us dividends in several other respects so that today even the western countries are getting interested in the functional attributes of Jaipur Foot. During the last few years, a whole series of new designs of footpieces is emerging in the west which are based on some of the design features of Jaipur Foot.

When we walk along a slope, our feet can turn in or out to adapt to the slope. In the old fashioned single axis metallic ankle joint of a western footpiece, this adaptation is not possible and so the entire artificial limb is deflected, causing considerable pressure on the skin of the stump at the stump-socket interface. In a more modern western footpiece, some degree of mobility is available and so the peak pressures at the stump-socket interface get reduced. But in Jaipur Foot, because of a much greater mobility, the stump gets very little pressure even when walking on uneven ground. In other words, our footpiece is comparable to the rubber bearings which are being talked about for earthquake-proof buildings. The base isolation by these bearings reduces the whiplash effect in which top storeys are literally shaken to destruction. These rubber bearings effectively "detune" the building from earthquake frequencies by a factor of ten. Our footpiece offers a similar protection.

A study was conducted in the University of Strathclyde at Glasgow where, in a sophisticated gait analysis laboratory, a comparative evaluation of the western foot and Jaipur Foot was carried out, using data recording the ground reactions through pressure transducers in a pylon dynamometer. The Scottish amputee who acted as an experimental subject was asked to return the Jaipur Foot after completion of this study. This amputee refused to part with the foot on grounds that it enabled him to



climb the mountains much more easily. Prof. Hughes, when presenting this paper, emphasized that this subjective, human response was far more valuable and revealing than any computerized study. This underscores the value of the weightage one ought to give to the user response.

Another interesting spin-off of our design was the capacity of the leg to rotate on the foot. The University of California group have been emphasizing that during normal human walking, various segments of our lower limb rotate on each other during different phases of the gait cycle. In the conventional western footpiece, the leg cannot rotate on the foot and this causes the entire artificial limb to twist around the stump while walking, causing considerable friction and discomfort. The Jaipur Foot, having got rid of the solid keel, allows this rotation and so the user of the limb is more comfortable.

The reason I am making this point is that while our original objectives were to provide a limb suitable for a floor-sitting culture, some of the spin-off is being held to be of basic importance even to our western counter-parts and there is now a renewed interest in the role of a footpiece as a dampener of ground reactions and what was formerly a low priority item on the research agenda in the west has suddenly been elevated to a higher level in the hierarchy of design of artificial lower limbs.

One can now match the list of our earlier objectives to what has been achieved. The foot fairly closely resembles a normal foot, and I often amuse myself by asking visiting surgeons to identify the amputated side. Even experienced orthopaedic surgeons have a 50% failure-rate! In fact, women often adorn their feet in a manner which has even fooled me.

The amputee can squat and one can witness the angle which the footpiece can make with the leg. There are amputees who are employed in our workshop who sit cross-legged on floor and work the whole day without the need for taking their limbs off.

The limb can now be used by our villagers, walking comfortably on a rugged terrain because of the adaptability of our footpiece.

The limb is waterproof and many amputees work in their farms, wading through water or mud. Drawing water for irrigation from a traditional well is a heavy duty job and yet these amputees perform such work like able bodied individuals. Rickshaw pulling is an urban vocation chosen by many poor amputees. They can even climb trees! The footpiece can grip the trunk and adapt to its contours. Such activities widen the

horizons of amputees who can continue to stay in their villages with their families and friends and carry out their former vocations. It is no longer necessary for most of them to migrate to urban areas, frequent the corridors of Social Welfare Ministry and end up with a sedentary occupation in an alien setting.

This is what 'true rehabilitation' ought to mean and it would be appreciated that there is a built-in element of rehabilitation in the design of these artificial limbs.

For the socket and the leg piece of our limbs, we opted for aluminium as a suitable material. Most of my colleagues adversely comment on the choice of aluminium. "The modern world is moving towards polymers and composites and you are moving back to metals", they comment. There are some very good reasons why I have preferred aluminium – atleast for below-knee limbs. We have skilled artisans in our country who can shape metal sheets with such ease and deftness that it takes one by surprise. A statue of a poor, emaciated amputee, which stands before the building of our Rehabilitation Centre was made by one of our craftsmen with aluminium sheets beaten into shape without any casting. It is a stunning piece of art. For people who have skills like this, and whose work adorn our handicraft emporia, shaping an aluminium limb is a child's play.

Visitors from abroad gape with amazement when, within 45 min., from start to finish, a below-knee trial limb is fitted. The tools for this work are simple; no plaster moulds are needed. The limb is shaped and fitted directly on the amputee who participates in the entire proceeding, guiding and informing the limb maker about the accuracy of the fit. This live human interaction between the amputee and the limb maker is a marvellous thing to watch. There is empathy and understanding between the two and a lot of feeling goes into this work.

Aluminium is available to us, easy to work with, light and strong and does not rust. Any pressure points can be easily lifted off with a tap of a mallet. Use modern FRP and you get into a much more expensive system where such manoeuverability after the resin is cured is just not available. So what is wrong about using aluminium? It is this simplification of the technology which enabled us to increase our turnover from one limb in a week in 1975 to ten limbs a day in 1982.

Materials are important but several considerations must go into their selection. I am not averse to new materials. In fact, we are the first in our country to use sophisticated materials like carbon fibre composites for rehabilitation aids and as of today, my choice for an above-knee amputee is a combination of flexible polypropylene socket with a carbon fibre load bearing frame. Having tried many materials, I find this, on various counts, to be a superior alternative. But for a standard below knee limb, my preference for aluminium stands.

Availability, cost, familiarity, physical properties, ease of modification, climate, skin allergy and many such factors must be put together and an entire range of options generated, from which an optimal selection ought to be made. What may be my choice at Jaipur may be different from what I might use in Nagaland or Bombay.

The approval of the design of our footpiece by the west has brought out another dilemma. There is now a continuous demand by centres from abroad for our footpieces. Jaipur foot centres are already functioning in Sri Lanka, Thailand, Indonesia, Peshawar and Zimbabwe. While this excites us at one level, it makes us very apprehensive at another level. Using rubber as our basic material, we use a very labour intensive technology. Our footpieces are heavier than the western analogues and they are not refined in their appearance as the western market would expect. There is a lack of standardization and no two footpieces have absolutely identical performance characteristics. We felt we should refine our product and turn to new materials such as polyurethane. The Department of Science and Technology came to our assistance and currently we are working on this material substitution. It has meant, of course, that the present properties of Jaipur Foot must be accurately characterized. The data base for the formulation of such a variable density polyurethane foot must be available before the polymer engineers can prepare proper formulations. This has forced us to generate this basic data with the help of structural engineers. This, I think, would be an extremely useful exercise which has not been carried out so far. At the same time, however, such materials require extremely critical operating conditions for manufacture. A much higher capital investment outlay is indeed, both for R&D and for setting up a production unit. If optimal conditions are not available, there can be a catastrophic failure of the footpiece. A rubber foot may not be as elegant but it is much less likely to fail. An analogy of the debate on traditional vs high yielding variety of wheat may not be out of place. Amulya Reddy is

fond of reciting the nursery rhyme — "When she was good, she was very very good, but when she was bad, she was horrid"

What we should not lose sight of is the "worst case scenario" rather than the "best case scenario" when evaluating costs and benefits. It is also important to resist the temptation of yielding to an applause from the west and in the process forget our rural amputees, for whom this work was taken up in the first instance.

Comment is often made, especially from the prestigious rehabilitation centres in our large metropolitan towns that our foot piece prevents their amputees from wearing fashionable shoes. "These are too broad", they say, "and there is a major problem of "foot entry" into narrow shoes". I concede this because my target group has been the barefoot walking rural amputee. Our footpiece has to match the broader splay foot of a barefoot villager. There is nothing to prevent, however, for another set of dies to be made for the urban rich, to produce footpieces which may have all the design advantages of Jaipur Foot and yet which can easily slip into elegant shoes. We have made a few such pieces, and with a detachable heel too, which can be inserted to preserve correct alignment of the limb, when the shoe is taken off at home.

There is a need, also, to update the technology for producing better rubber Jaipur feet. A lot of progress has been made in rubber technology since we started working in this field in 1965. Better rubber formulations, lighter and stronger, improved die design and a production technology which can ensure greater standardization and quality control checks, should be effective in overcoming some of the existing shortcomings of Jaipur Foot.

There are thus several options available which could be pursued:

- (i) The Jaipur Foot could exist as it is; it is inexpensive and has stood the test run in atleast 20,000 amputees.
- (ii) For urban rich, a modification could be used to allow easy foot entry in fashionable shoes.
- (iii) Improved rubber feet could be produced using updated rubber technology, or
- (iv) Polyurethane could be substituted as a better material for export quality footpieces. Depending on the nature of consumer demand, a

footpiece could be made available for different population groups. All these options could exist side by side.

### APPLIANCE FOR POLIOMYELITIS

The second example I would want to present deals with the polio problem. Poliomyelitis is the largest single cause of physical disability in our country. With its disappearance from the west, no new ideas are coming forth from abroad on designs of appliances for poliomyelitis. Whatever new thinking had to go into this problem has now to be generated by us.

It is not commonly appreciated that it is much more difficult to design appliances for polio than an artificial limb. This is because the polio child has a choice. If he does not walk better with the device, he just won't use it. The amputee, on the other hand, has no choice. One of the major problems in polio is when the muscle of the thigh, the quadriceps, is paralysed. The knee joint then becomes unstable and the child has to use his hand on the thigh and press the knee backwards to prevent the knee from buckling. There is no suitable operation which can stabilize the knee and so a metal brace is used to lock the limb into it. This liberates the hand and now the child can walk upright without the fear of falling down.

Such metal calipers have been in use, without a major design change over the last century. While we keep on prescribing them, any honest follow up would reveal that there is a very high rejection rate by the users. The reasons are not difficult to understand. These calipers are heavy and the already paralysed limb has to drag this extra weight. The knee cannot bend while walking and this poses a problem of clearing the ground while swinging the limb forward. The lower limb, which normally behaves like a compound pendulum, is converted into a simple pendulum with a long lever arm. This demands an extra effort by the muscles to swing the limb. And so, instead of acting as an energy-saving device, this caliper in fact becomes an energy-consuming gadget. Since the metal side bars are fixed to a heavy duty shoe which is the foundation of the device, the child gradually starts nursing a hostility towards these shoes which are not interchangeable and different from those worn by other children. It should not surprise us, therefore, that our patients usually reject this solution.

This basic design of a caliper is related to the use of metal side bars. Metals can only be used in certain ways and impose a tyranny of their own. But now a variety of new materials are available which can be easily shaped and weight for weight, some of them are stronger than metals.

Thus, we can now think of different ways to stabilize an unstable knee by using new geometries of designs. We have now been using a carbon fibre composite, a spin-off from aero-space engineering, and utilizing a totally different design concept. By artificially keeping the foot at an angle so that instead of the heel, the front part of the foot strikes the ground first, and with the help of two lateral bars and a cross piece in front of the knee, the device behaves like a cranked lever where the body weight from above is used to push the knee back and prevent it from buckling.

This so-called floor-reaction orthosis has now been used in over 500 cases with an 85% acceptance rate. It is four times lighter than the metal caliper, allows the bending of knee while walking and can be used with any shoes. Carbon-fibre has poor abrasion resistance but a polyurethane sole, stuck to it, can probably allow our rural patients to walk barefoot in their farms.

The cost of this appliance turns out to be actually cheaper than a conventional caliper. The technique of fabrication does not require any expensive outlay and anyone who enjoys working with his hands, without any formal education, can be taught to make this appliance. It still suffers from disadvantage that carbon fibre is an imported material, the selection of cases suitable for this appliance requires considerable experience, each appliance has to be custom made and the margins of error permitted are very narrow.

This is only a beginning. Already a number of new designs are under trial and we foresee a major shift from metals to polymers in these appliances in the near future.

Because of the need for shoes, a major problem in the logistic of supply of conventional calipers is encountered. Not only are these expensive but they need to be custom made. An analysis of the reasons for long waiting lists in the delivery of calipers in most rehabilitation centres reveals that the major hold-up occurs in the footwear section. Recalling that closed shoes are not well tolerated in our warm climate, and to speed up production, we had worked on the idea of substituting shoes with wooden clogs, an idea initially mooted by Huckstep in Uganda. Huckstep's design of a wooden sandal was too simplistic and inefficient. So we had to work out a new design, with a roll characteristic which allowed a much better gait. This wooden clog could be made very quickly and with prefabricated leather straps and a large stock of such clogs of different sizes, it is feasible to fit the child on the same day that he is seen.

The design of a caliper with its wooden clog was so simple that the idea of using our village craftsmen to prepare such calipers occurred to us. Every village has a cobbler, a carpenter and a blacksmith. They are needed by the rural community. Why can't we use this available manpower so that polio children could be fitted with braces in their own village? An experiment was conducted at Tilonia, through the help of Bunker Roy. We took our caliper there and showed it to these craftsmen. They were able not only to make these with the locally available materials and tools but actually came out with a product which was superior to ours in workmanship. The innovative capacity they demonstrated in making their own version of a limited motion stop at ankle came as a surprise to me. Who says our rural poor are not intelligent?

This provides a totally new dimension to a problem which is confounding our administrators sitting in Delhi, counting the millions of polio children spread out over our vast country and wondering how to handle this massive problem. Looked at a village level, however, where there may not be more than 4 or 5 cases per village, it can be a very solvable problem provided we are willing to leave our institutional hideouts, share this knowledge with village craftsmen and encourage them to become self-reliant in providing a neighbourhood facility for preparing atleast simple rehabilitation aids. I know how our professionals react to such ideas which they find outrageous. This is truly a Hobson's choice. On the one hand such a strategy may provide *inadequate aids*. On the other hand, adequate technologies are *inaccessible*. Sticking to the idea of providing only the best usually means that 90% of our poor population have to go without any aid whatever.

There are thus three different kinds of options available for aids in poliomyelitis. One is to continue with the present strategy of a centralised production agency, supplying factory made metal components to be assembled and fitted locally. Any change in design is not permitted here and one needs a large bureaucratic machinery to manage the supply and fitting.

The other is to continue research into the use of new material and new designs and field test them. A lot of R&D effort has to go into this option because it is now futile to look to the west for any fresh ideas on poliomyelitis.

The third is to simplify the existing designs and work out a strategy of using rural craftsmen to provide a neighbourhood facility.

The costs and benefits of each will have to be worked out but in all this exercise, let us not forget the user – who would usually belong to the group of rural poor. The constraints of time prevent me from multiplying such examples endlessly.

The main point I have tried to make is that in a dual society such as ours, and this is true of all developing countries, we are constantly running into a Hobson's choice. The technologies and designs evolved in the west are preferred by our rich urban elite and they really constitute the market forces which influence our bureaucratic machinery. The poor are outside the market forces and have no voice. Modern technologies are inaccessible to them. To permit the poor to escape from this dilemma, scientists and technologists must generate new options, each more effective than the traditional, and more accessible than the modern. Ideally, the options should constitute a hierarchy of technologies with upward compatibility. Then, with rising incomes, the poor can climb from a cheaper, less cost-effective option to a costlier, more cost effective option. Only in such a situation will the people have genuine choices. Thus, the role of scientists and technologists is to be option-generators and choice-wideners.

People who control decision-making in our country are understandably in a hurry. They overlook that a more appropriate and equitable generation of technology involves a "learning curve". During the initial part of this learning curve, there has to be intense back-and-forth interaction between the lab and the field. The feedback from users in the field must lead to modifications and improvements of the product/process. This modified/improved product/process needs further "test marketing" in the field. As a result of this interplay between technology generation and dissemination, and between technologists and potential consumers of the technology, the penetration of the "market" is necessarily very slow during this phase. Only later, our learning curve shows a steep climb.

All these points are generally ignored when technology dissemination is planned and implemented. There is a general tendency for technology generation and technology dissemination to be thought of as two distinct non-overlapping sequential stages with the generation ending when the dissemination begins, and the generators "washing their hands off" the technology dissemination process.

However idealistic and romantic it may appear, my conviction is that the technologists must approach such work with empathy and affection for the people. Otherwise, they tend to be afraid of the people



and hide behind their institutional walls. The poor are far more understanding of our failures than so-called educated who cheer when the satellite goes into the orbit and jeer when it falls into the sea.

Science and technology ought not to be "value-free" and would stand to gain from these feelings of empathy and affection. Without this value-system, it tends to become amoral, unjust and violent.

A lot of hard and painstaking work lies ahead of us. The problems facing us are open-ended. This is why I am worried about a 'top-down' managerial approach which, some people think, will quickly solve our problems. Bernard shaw's approval of "the inevitability of gradualness", carries for me, a lot of wisdom.

**Vergheese Kurien** (b. 26 November 1921) did M.Sc. (1948) from Michigan State University; He was awarded Doctorates (h.c) from several universities. He is Hony Chairman, Gujarat Co-operative Milk Marketing Federation, Anand (1983- ) and National Dairy Development Board, Anand (1965- )

Kurien's contribution to Dairy Development have been internationally recognised. He played an important role in bringing the white revolution to India, which was started from Cooperative patterns and created marketing facilities for milk and milk products. He also helped in improving the Indian cattle by advocating cross breeding with foreign blood.

Kurien is Fellow of Institution of Engineers (India); Patron, Indian Dairy Association, New Delhi; Hony Member, Association of Food Technologists (India), Mysore, Companion, British Institute of Management, London, UK. He is the recipient of several awards; some of them being: Wateler Peace Prize by his Royal Highness Prince Claus of The Netherlands (Carnegie Foundation), Hari-om-Ashram Prerit Prof JP Trivedi Award (1987); Prof JG Kane Memorial Award (Oil Technologists Association of India) (1987); Silver Jubilee Award (Indian Society of Agricultural Engineers) (1987); Shiromani Award, Delhi (1988); National Integration Award (Indian Chamber of Commerce) (1989); World Food Prize Award (General Foods Fund, Inc., USA) (1989); Professor B.D. Tilak Lecture Award (INSA) (1990); Padmashree (1965), Padmabhushan (1966).

**Address:** Hony Chairman, National Dairy Development Board, Anand 388 001;  
**Tel:** 22041 (Off) 23000 (Res); **Telex:** 0172/207.

## **FOOD AID AND INTERNATIONAL AGRICULTURAL DEVELOPMENT: THE AGENDA FOR THE 1990s**

V KURIEN

When I first heard from Professor M M Sharma, the President, INSA, that the Indian National Science Academy proposed to honour me by conferring on me the B D Tilak Lecture Award, I was delighted. However, I must admit that my happiness was tinged with apprehension. I have wondered since if I really deserve the honour. Even today, I must confess that I feel somewhat dwarfed in the presence of such a distinguished gathering of scientists and scholars.

For, I am not a scientist myself although I had wanted to become one 40 years ago. As a matter of fact, I was well on my way to becoming one, and a pretty good one at that. But life has its secret ways of nudging one in unforeseen directions. And soon after returning from abroad with American degrees, I found myself mixed up with the farmers of Kheda. My scientific career was thus summarily terminated; and instead of spending my life in laboratories and classrooms, I ended up working for farmers.

I have never regretted this. And, to be frank, given a choice, I would not trade one bit of a life devoted to the service of farmers for a lifetime as a scientist. In point of fact, I believe that I have been able to do more for science as well as for development by helping farmers to create organisational structures that put modern science and technology—the main instruments of development—in the hands of the people.

Therefore, it is that, when I received this invitation from Professor Usha Luthra, I was wondering what could a man like me say to a group of eminent scientists which they may find significant. After much debate, I decided for the wiser course: rather than talking about science on which most of you know more than I do, I might talk on the linkages between food aid and economic development, a subject which I perhaps know more than most of you do.

## THE RISE OF AN 'INDIAN' DAIRY MOVEMENT

In the years after India's independence, the government of India experimented with a number of different approaches to developing our national dairy industry. There were milk colonies, government dairies, intensive cattle development programmes, key village schemes. It seemed, however, that all these programmes succeeded in increasing the number of bureaucrats employed to run them. As time passed, or milk production stagnated and per capita consumption of milk began to decline at a precipitous rate.

On this rather bleak landscape there was one notable exception: this was a cooperative enterprise started by some men of vision, ability and principle, an enterprise that linked the tremendous productive genius of our farmers to professional management and modern technology. It was my good fortune to be associated with this enterprise from an early stage.

In 1964, the then Prime Minister of India, Shri Lal Bahadur Shastri, visited this enterprise in the small town of Anand in Gujarat. He spoke with the farmers, looked at their fields and at their animals. What he concluded was that what had made Anand a success was the fact that the farmers themselves had control over their cooperative and that they had the wisdom to employ component professionals to manage that enterprise for them. He, therefore, decided that we should try to create Anands throughout India. To this end the National Dairy Development Board of India was created.

When we started out in 1965, we had a reasonably clear idea of how to proceed with the creation of Anands. But we had few resources with which to work. It soon became clear that the bureaucrats and technocrats who controlled all those government dairies, milk schemes, cattle development programmes and veterinary departments were not at all anxious to surrender their empires to work for a bunch of farmers. At the same time, we were becoming increasingly concerned as we watched mountains of butter and skim milk powder growing in Europe. It would only be a matter of time before those mountains were exported to India. Had that happened it would have tolled the death knell for our nascent dairy industry.

Fortunately, some men of vision collaborated with us to transform those mountains from a threat to an asset: initially with the support of the World Food Programme and then directly with the European Economic

Community we were able to use food aid as an investment in the revival of dairying in India.

That investment of food aid, later combined with World Bank assistance, has produced results. Today, more than six million families participate in a dairy cooperative structure that covers much of India. Milk production has increased substantially as has our per capita consumption of milk and milk products. In 1989, our production had reached a level where we would decline dairy commodities as food aid; we shall again decline such assistance this year. Save in the event of severe natural calamity, we hope that we have now reached the point where we shall no longer require such donations. In fact, I suppose that it is no secret that this year we have made a modest entry in the export market.

More important than increasing the availability of milk, 'Operation Flood' has meant a reliable source of income for millions of our rural people. For many, dairying has become their primary employment. In fact, the success of the cooperatives has increased employment in every area of the dairy industry, from equipment manufacture to marketing of products. I believe today that milk stands second only to rice as a source of agricultural-based income in India.

The success of 'Operation Flood' in dairying has led to its adaptation and application to oilseeds, horticulture and even forestry. Although these are still in a relatively early stage, the results are also quite encouraging. In fact, all our experience has convinced us that food aid can and should be used as an investment.

### FOOD AID AS AN INVESTMENT

Donors generally approach food aid with a mixture of interests and motives. Certainly there are humanitarian concerns involved—the desire to alleviate the misery and suffering of one's fellowmen, a motive that may also reflect a sense of guilt over having contributed to creating the conditions in which those fellowmen suffer. Donors also use food aid to bolster their political relationships with the recipient; as a means of disposing the surpluses created by their own farmers; and as a strategy to develop future markets. It is hoped that today's recipient of largesse may become tomorrow's buyer of the same product. I think you will find all of these interests incorporated in the legislation that various donor countries have enacted on the subject of food aid.

Recipients, too, approach food aid with a variety of motivations. In some instances, faced by natural or man-made crises, food aid is requested in order to save the lives of those affected. In other instances abundant supplies of food aid mask the failure of government's agricultural policies and help to constrain what might otherwise become very high food prices—particularly in the urban areas—which seem to be the focus of concern for most governments; quite possibly because politicians and bureaucrats generally live in cities. Another motive in seeking and accepting food aid is to provide nutrition to what we might call the weaker sections of a population, those urban and rural poor who cannot afford an adequate and balanced diet.

These various donor and recipient interests tend to be expressed primarily in two types of food aid programmes: emergency assistance and humanitarian relief.

In the case of emergency relief, food may be a vital necessity to stave off the results of natural disasters such as droughts, floods, earthquakes, cyclones and the like. In countries which live on the thin edge of survival, such assistance can prove the difference between life and death for thousands, or even millions, of those affected.

Food aid may also have practical and humanitarian importance in the medium and long-term alleviation of hunger and malnutrition. Here, too, it may be vital to the well being of its beneficiaries.

However, laudable as such programmes may appear, and as politically defensible as they are for the donor governments concerned, there would appear to be mounting evidence that the results are not all that is desired—whether viewed from the perspective of the donor or the recipient.

First, we have seen that large supplies of food, whether as emergency relief or humanitarian assistance, can often trigger a cycle of dependency. Because the food is readily available—both to the hungry and to the governments who are responsible for the hungry—inadequate efforts are made to identify and correct the problems that require external assistance in the first place. I am told that today in India—and I expect in other parts of the world as well—the children and even the grand children of the original "beneficiaries" of feeding programmes continue to rely on such donations.

Second, the easy availability of free or heavily subsidised food will depress domestic prices to the point where it is no longer feasible for the local farmer to compete. As you can imagine, this simply exacerbates the problem that triggered the need for food aid in the first place.

Third, we are told that there are instances where consumer preferences for donated cereals and processed foods have displaced local foods from the diet, again at the cost of local self-reliance.

I believe that there have been any number of studies done by economists and social scientists (may be the tribe perish— they are never there, where the action is), studies that have highlighted these types of problems with the use of food aid for emergency relief and humanitarian assistance. Despite these, the lion's share of such assistance continues to be directed to these uses.

There is a third use of food aid: food as an investment in the production, processing and marketing of food. In this way, I suggested, food could be used to eliminate the need for such aid in the future.

However I must add a note of caution. Using commodities as an investment in development is not a panacea. Our experience has made it clear that there are a number of important conditions that must be observed if such an approach is to prove a success.

The first condition is that we create producer-owned and controlled structures. These enterprises return a far greater share of the consumer price to the farmer. They build markets, supply inputs and create value-added processing. All this happens because the farmers' productive capacity is linked with professional management in cooperative organizations that have staked out a place in the market. Put bluntly, these structures force others in the dairy business to compete fairly and they turn the terms of trade in favour of the rural producer.

When producers have such structures at their command, we know that they have the means and the will—to ensure that the results of science and technology reach all those who will benefit. It is only when such structures exist that farmers gain the confidence necessary to stimulate their investment to increased productivity. It is only when such structures exist that farmers can demand—not beg for—the services and inputs they need to realise returns on that investment.

Let me share the greatest lesson we have learned: we must respect and trust our farmers. They may not be educated, or even literate but they

possess uncommon sense and even uncommon wisdom. Time and again they have shown the ability to rise above narrow self-interest to act together in pursuit of a greater good. Programmes and projects that are designed and managed from the top down—that call farmers "targets" – and that ignore the untapped capacity of rural producers to manage their own affairs, such programmes invariably fall short of their goals. Those that respect the right of the farmer to manage his own affairs cannot fail.

It is not enough, however, to use food aid as an investment and to promote farmer-controlled enterprise. Such structures must operate in a social, economic and political environment, one that may extend well beyond national boundaries. Therefore, a concerted effort must be made to ensure that this environment includes the programmes and policies that encourage and support development. Pricing, import and export regulations, taxation, investment terms and conditions are among the many such policies that must be coordinated in an overall approach in support of food aid.

The donor, too, has a role to play in ensuring the success of food aid as an investment.

The first point concerns quality. Unlike emergency relief and humanitarian programmes, food aid as an investment requires that consumers pay for the food. In order to ensure that local prices are not depressed, the consumer must pay as much as he would if the food was produced locally. This means that the quality of the commodity must justify such prices: in other words, food aid for investment must be of the finest quality.

Second, building an industry from basic infrastructure to production through marketing takes time. It takes even longer if a strong producer-owned structure is to serve as the foundation. This requires donors to think in terms of time frames quite different from those to which they are accustomed. A programme of Operation Flood's magnitude may require two decades; others, somewhat less ambitious, may require only seven, ten or twelve years. What is critical is the continuity of the support. Once commodity assistance is provided, it must be committed for the duration. This means that food aid as an investment must be insulated from short-term problems of supply and from political exigencies.



## FOOD AID AS DEVELOPMENT INVESTMENT : CONSTRAINTS

Twelve years ago I suggested that the overriding objective of all aid should be to eliminate the need for aid and that the use of food aid as an investment would seem to be the most likely way to achieve that objective. I then asked, "Why is it that food aid has rather seldom been used as an investment?"

I wish that, today, I could say that food aid is now widely used as an investment. Regrettably, I cannot. While we have successfully used food aid as an investment for Operation Flood, and with other commodities, there would seem to be few other examples of such uses of food aid elsewhere in the world, or even in India. This is despite the fact that each year we receive visitors from every corner of the world, visitors who want to attempt what we have achieved in Anand. Each year we receive invitations from these countries and, to the extent we can, we send our officers to provide advice. But, without food as a resource, little can be achieved. So, in 1990, I must again ask why food aid is not widely used as an investment?

One explanation that comes to mind is that our very success in using food aid may have caused donors to have second thoughts. Several donor countries not only explicitly include market development as a rationale in their food aid legislation, they have also placed restrictions on economic and commodity assistance that might help the recipient to become self-reliant or, heavens forbid, a competitor. So I must ask myself the question: has India by making a success of food aid—that is, by eliminating the need for such aid—prejudiced the chances of other countries to do the same?

A second explanation that has been advanced is that the great mountains of butter and skim milk powder are no longer there. Government programmes to buy herds, to limit acreage, to penalise productivity, have all combined to bring supply into close balance with demand, sufficiently close to end surpluses and, if you will, nudge prices up a bit. While there is no denying that such programmes have been implemented, there is apparently also no denying the will of the farmer to produce. And so, I am told, once again mountains of butter and powder are beginning to rise.

A third explanation is that foreign assistance is never popular with domestic political constituencies and that is far easier to convince a skeptical public of the need to support programmes to feed starving people

than to explain and defend the intricacies of food as an investment. There are few politicians willing to honestly admit to the fact that all official development assistance represents only a small fraction of total resource transfers between rich and poor nations and that the flow of those transfers continues to move from poor to rich.

### AN AGENDA FOR THE 1990's

Whatever be the reasons, I believe that we must take a broader and longer view. We can no longer afford to do otherwise.

We have enjoyed a decade of "almost enough". Just enough to cause complacency. If there has been a bit of hunger in the Sahel, then we have patched it up by holding a few concerts and sending a few shiploads of food.

We are now looking at a far more difficult decade, one that will cause many a rude awakening. The world faces a grave agenda; poverty, hunger, re-emergence of oil as a source of crisis, deteriorating environments, growing populations, new and dreadful diseases.

There is hunger in today's world. If India is self-sufficient in cereals, it is because a significant portion of our population can afford to buy very little. The same must be true of many other countries. Even in the industrial nations the hungry exist in growing numbers.

Yet, in the face of hunger we see the governments of some of our most productive nations adopting policies and programmes to reduce production of food. Commodities available for donation, and even for sale, have diminished. This has not only reduced world supplies, it has caused hardship to the producers and consumers of the very countries that have adopted such policies.

What must be the effect on the farmer who is told not to achieve excellence, who is bribed to do less than his best? Knowing, as I do, the enormous benefits that result when food is used as an investment, I feel all the more strongly that to restrict production in a world with hunger is wrong—even immoral.

And, I might add, it is ill-considered and short-sighted.

Agriculture is important to the economies of even the most advanced nations. The exports of agricultural products is critical to the health of the economies of countries like Canada and Australia. Many of

these nations have "invested" a great deal in building commercial markets through commodity assistance. Yet, by creating and reinforcing dependency, they have ensured that the recipients of their support remain unable to afford the products they wish to sell.

To compound the problem, we see national governments spending vast sums to take land and animals out of production. This has driven up the price of their commodities. So they spend even greater sums to bring the price down to compete with other exporters who have done precisely the same things. And the next thing we know the IMF and World Bank start talking about "comparative advantage" and telling us to liberalise and rationalise our economy! Comparative advantage, it seems, is having the resources to subsidise producer and export prices.

So we have a situation today where the poor nations of the world are gripped in a cycle of poverty and debt which is unwittingly reinforced by the policies and programmes of the wealthy. But the wealthy find that in order to sell to the poor, they have to either subsidise the price or simply donate more—either action simply reinforcing the same cycle of poverty with the only difference being that each year the number of poor grows, increasing the misery and raising the cost of its alleviation. Where does it all end?

Today the per capita consumption of butter in Europe and the Soviet Union is more than four kilograms annually. Pakistan's citizens consume just under three kilograms per head; Canadians and Americans consume around two kilograms; we Indians divide our butter up at the rate of about one kilogram for each of our 844 million. The rest of the world consumes, on an average, about 340 grams per capita annually.

I can only assume that if the 3.3 billion people who consume that 340 grams could afford to do so, they might wish to use a bite more butter—let's say a kilogram each? To meet that increased demand would require a little less than 2.2 million tons of butter, almost double the current world trade in the commodity. Instead of squabbling over who gets what share of today's market, should the nations of the world not be finding ways to ensure that tomorrow's market is far larger? The "surplus" of world dairy commodities is only a surplus as long as much of the world's population can't afford those products. It is, therefore, in everyone's interest that they be able to afford them, and as rapidly as possible. By using today's agricultural surplus as food aid for investment, it should be possible to help build the agricultural economies of those vast

sections of our world that are poor, advancing the day when they become buyers instead of borrowers or beggars.

Unfortunately today there are ominous clouds on the horizon. There are those who believe that the rapprochement between East and West signals a decision to realign the North against the South. In a world with finite resources and growing populations, those of short-sighted vision may try to erect a fortress to protect their disproportionate share of the world's wealth, a share that was often garnered through exploitation and the expropriation of other nations and peoples.

Such a confrontation between North and South would be both tragic and unnecessary. Particularly when we still possess sufficient time and resources to rebuild the economies of the world's poor nations, helping them to become self-reliant partners in a far stronger world economy. Food aid, as an investment, should be a major strategy in that effort.

It is said that the dairymen of the industrial nations are very effective lobbyists. How else to explain how a handful of men and women have managed to ensure that their governments protect and advance their interests, often at substantial costs to the national exchequer? Let me then propose that these skills be employed on behalf of the world's poor, not simply to alleviate their suffering, but to enable them to become productive members of a stronger world economy, a world economy where production need not be restricted because three billion consumers cannot afford to buy what they need.

Food aid has played a very important, indeed a critical role in the development of India's dairy industry. Food aid can and should play an equally important role in building the dairy industries of other Asian countries, in Africa and Latin America. During the generation that it will take to achieve this goal, this food aid will offer an outlet for the productive capacity of the major dairying nations. It will provide the time, if needed, for a gradual, orderly and rational restructuring of those nations' dairy industries— not the brutal, chaotic and immoral measures we have seen in the last few years. The costs will be marginal; the returns beyond measure. This, then should be the agenda for the nations of the world as we approach the 21st century.

Let our goal be the ability of each nation to feed itself. By that I don't mean that each nation needs to produce all that it consumes, but that each will produce enough to ensure that what is not available can be

purchased. Nor do I mean that a nation is feeding itself when large segments of its population can afford only the barest minimum. A nation is feeding itself when the nutritional requirements of all its men, women and children—especially its children—are fully met.

Let those among us that have— or which can produce a surplus—sit together with those that are as yet unable to fully feed themselves, and let us evolve a plan in which surplus food will be used as an investment to increase the production and productivity of those presently in need. Let us plan, produce and channel surplus to meet the requirements of countries in need—not haphazardly dispose off the occasional excess.

Let us plan and commit resources over the span of a generation, sufficient time to ensure genuinely sustainable results.

Let us jointly examine and agree upon the types of policies and programmes that serve a world in need, not the interests of a few. Let us agree that food aid investment serves the interests of both donors and recipients and commit ourselves to the highest standards of quality, continuity and integrity—not only in commodities but in all our efforts.

In the past, through support of Operation Flood, the world's dairy industry has pioneered the use of food as investment. Let that pioneering step be but the first of many in which the world's dairymen show the world the way to use the fruits of our labour to benefit mankind. On our part, we in India are prepared to commit our experience, our human resources, and to the extent we can, our own commodities, to achieve this goal. This is the way we would like to salute those who helped us herald the white revolution in India.

**Noshir Hormasji Antia** (b. 8 February 1922) did MBBS (1945) from Grant Medical College Bombay and FRCS (1952) England. He is Director & Trustee, The Foundation for Medical Research; Director & Trustee, The Foundation for Research in Community Health; Hony Consultant, The Tata Department of Plastic Surgery and J.J. Group of Hospitals; Recognized Teacher in Applied Biology, University of Bombay; Hony Surgeon to the President of India; Hony Surgeon to the Governor of Maharashtra.

Antia established the plastic surgery department in JJ Group of Hospitals and Grant Medical College, Bombay and performed the first clinical case of microvascular surgery. He has developed special sections for surgical treatment of burns and leprosy, and established rehabilitation centres for those deformed by leprosy, burns and other physical disabilities. Established, for the first time, department for external facial prosthesis, using the latest silicon technology. He established the Foundation for Medical Research, for advanced research in neurology, immunology and microbiology of leprosy where *Mycobacterium leprae* was cultivated in vitro within the schwann cell, for the first time. He has also studied the role of macrophages in leprosy and showed the importance of this disease which has helped in screening of drugs and study of drug resistance in leprosy. Further, he has done much work on community health, problems of health care and health services.

Antia is Fellow of National Academy of Medical Sciences; Member (Past President) Association of Plastic Surgeon of India; Indian Association of Leprologists; President, International Confederation of Plastic and Reconstructive Surgery; Senior Member, British Association of Plastic Surgeons; The International Society for Burn Injuries; Hony Member, Indian Society for Surgery of Hand; Japan Society of Plastic & Reconstructive Surgery; Member, Association of Surgeons of India, British Society for Surgery of Hand; Retired Corresponding Member, American Society of Plastic and Reconstructive Surgeons and its Education Foundation. He is the recipient of Padma Shri (1990); Menino D'Souza Oration Award (1984); K.C. Sahu Gold Medal (Hind Kusht Nivaran Sangh) (1979) and Professor B.D. Tilak Lecture Award (INSA) (1992).

**Address:** The Foundation for Medical Research 84-A, R.G. Thadani Marg, Worli, Bombay-400 018, India; Tel: 4934989, 4932876.

## **NEW DIMENSIONS TO HEALTH CARE**

N H ANTIA

Bal Dattatraya Tilak is one of the rare breed of scientists who can differentiate between the myths of western science and technology and its reality. This has enabled him to discern the difference between knowledge and wisdom, between the glossy packaging and the contents, and between sophistication and elegance. This is why he has been able to use his scientific training and talent to solve some of the major problems that beset our people who live in rural India where the heart and soul of 'Bharat' still resides.

I met Dr Tilak for the first time in a small basement laboratory at the Indian Institute of Education, an institution built by another great son of India, the late J P Naik, where the previous director of the National Chemical Laboratory was dirtying his hands with what looked like crude chulas to be used by the rural poor. The significance of what he was attempting to achieve struck a common chord, for here was one of our eminent scientists turning his talent to solve a major problem of our country which, though without the glamour of high technology, had the potential of easing the life of millions of women and children from gathering of wood and saving their health from the acrid fumes of traditional cooking in their huts. That this simplified technology could also be a major contribution to the saving of both human and fossil energy, of preserving our vanishing forests and may be towards world ecology. In this, Professor Tilak was attempting to blend modern science with our traditional culture and establish a symbiosis of both, and in the process not afraid of swimming against the tide.

I shall try in this lecture devoted to him to demonstrate how the same principles can be utilized in the field of health and medical care.

India has one of the world's oldest and richest traditions in the field of medicine and health care. Two thousand five hundred years ago there already existed a well-developed ayurvedic system for medicine and surgery. And yet today the system which exerts the greatest influence on the present medical scene is allopathy introduced by the British primarily for the care of their own expatriates. It is an approach which has proved highly effective in elucidating the germ theory for communicable diseases

as also in the field of surgery. However, with its focus on detail, allopathy has its inbuilt limitations which have prevented the development of a holistic approach to health, the very essence of Ayurveda and Yoga. These indigenous systems that emphasize maintenance of health rather than mere cure of illness do not isolate the mind and spirit from the physical body and the human being from its natural environment. The shortcomings of the allopathic system is not only because of the narrow dissective nature of this science but also because it carries with it the aggressive and exploitative culture from which it emanates. While providing knowledge and technology which could meet the basic needs of the entire human race and free it from drudgery and slavery of the past, western science and technology has also provided its masters with endless material goods as well as vast military power. This has been used to colonize other countries whose rich natural resources they could appropriate for their ever-expanding industries. Devoid of spiritual content or moral restraint this new science and technology suffers from this 'original sin' being born as a reaction to the dogma of the Christian church in the middle ages. Arrogance of power and the absence of holistic understanding have prevented utilization of the beneficial aspects of this science and technology which now poses a threat not only to those who do not have access to it but also to the fauna and flora which sustains all life on this planet. Its pandering to human greed, which knows no limits, explains why modern medicine has also been converted into an unnecessarily expensive and lucrative business based on exploitation of human suffering while neglecting its most beneficial role in the realm of communicable diseases.

At Independence our founder fathers rightly decided that the benefits of health care, hitherto restricted to a select few, should be made available to all our people. For this we were fortunately provided in the Bhore Committee's Report of 1946 a detailed report on the health of our people as well as a blueprint for devising a service for the entirely new requirements. This was the original people-based Primary Health Care concept and was adopted for the post Independence development of our country's health services. This still remains the most farsighted document for Health for All and has been reiterated by the WHO at Alma Ata in 1977, the ICSSR/ICMR report in 1981 and incorporated in our National Health Policy of 1983.

The development of health services and health care of our people in the four decades following Independence can be best understood if studied



separately in the first two and the last two decades. Let us now examine what happened in the first period following our independence when there was a strong political will to make available the benefits of health care as of all development, to the entire people. The Bhore model was accepted as a natural blueprint for the health services of post-independence India. It was suited not only because of its appropriateness but also because there already existed a readymade formalized system consisting of doctors, medical colleges and hospitals; also because it had already demonstrated its effectiveness in controlling malaria, smallpox, water-borne diseases, cholera and plague in the British cantonments utilizing the knowledge of the cause and spread of these diseases. Even in the absence of DDT and antibiotics, the existing knowledge of preventive medicine could prove to be highly effective when used by an efficient public health system with trained and disciplined para-medical personnel functioning under the supervision and guidance of the IMS officers. This nucleus now had to be expanded to cover the entire country requiring massive extension and deployment. Given a strong political will together with full support and political non-interference there was a remarkable demonstration in the early post-independence era of what could be achieved by the use of this aspect of Western science and technology. Malaria whose prevalence rate was in the region of 75 million cases per year was reduced to 65,000 by 1965 and the 2 million annual death cause by this disease virtually eliminated. Vast highly endemic regions like the Terai were opened up for human habitation and development. The availability of DDT and synthetic quinines certainly helped, but more important was their systematic utilization, even in remote villages, by a large team of paramedicals under the supervision of a small but devoted cadre of public health personnel trained in epidemiology and managerial techniques. Small pox for which a vaccine was available for over a century, but which had continued to pose a major health hazard, was eliminated due to effective use of this tool. Cholera and plague the other two major health hazards of this period were also brought under control.

This remarkable demonstration of the effectiveness of public health measures for communicable diseases on such a massive scale in a period, when the rest of the health infrastructure was limited, remains a saga in the history of health care of our country which has not been repeated to date despite the availability of far more effective measures in the form of insecticides, vaccines and drugs and a vastly expanded health manpower and infrastructure in the public, private and NGO sectors.

Table I

*Health infrastructure*

|         | 1956     | 1988     |
|---------|----------|----------|
| Beds    | 1,52,888 | 6,09,735 |
| PHCs    | 725      | 16,756   |
| Doctors | 61,440   | 3,52,000 |

This also resulted in an increase of life span from 41 to 59, reduction in the death rate from 19 to 11, birth rate from 43 to 31 and of IMR from 146 to 94. Improvements in nutrition as a result of increase in food production and its distribution also played a role in improving the health status of our people during this period. The allocation of 5% of the budget to the public health sector in the First Five-Year Plan demonstrated the importance paid to the health of our people which has not been repeated in subsequent Plans.

There has been a radical change in the health scenario in the following two decades. Despite massive increase in health inputs in both the public and private sectors, diseases like tuberculosis, leprosy, tetanus, gastro-enteritis and acute respiratory infections continue to take their unremitting toll notwithstanding the availability of improved knowledge and technology for their prevention and cure. Even malaria which had been virtually eradicated showed a resurgence despite the continuation of a vertical programme which consumes a major portion of our limited public health budget. By this time the emphasis had shifted from preventive to curative medicine and that too for diseases like cancer, heart and stroke which mostly affect the urban rich, rather than the far more common communicable diseases which continue to plague the rural poor. Those trained in Government medical colleges preferred to emigrate to greener pastures or work in the for-profit private sector. It was hoped that these doctors trained at public cost would provide leadership to the large team of para-medical workers in Primary Health Centres designed to provide preventive, promotive and basic curative service to the rural people which still comprises 70 per cent of our population, the majority of whom cannot afford private care except under duress. Table II demonstrates this disparity in the rural-urban distribution.

For this they were neither appropriately trained for this task nor were they willing to serve in the rural areas. Like the rest of the elite they

have pursued their own self-serving goals. The lucrative nature of the every expanding private sector is demonstrated by the exorbitant capitation fees paid in declared and undeclared monies by the sons and daughters of the rich to the politicians who promote such institutions purely for personal profit under the guise of health care for the poor. The professional-bureaucratic-political nexus has also been responsible for permitting the production of Rs. 3,800 crores of drugs by the pharmaceutical industry, the majority of which are unnecessary and some even banned in their mother countries. Instead of the 250 essential drugs recommended by the WHO and 2000 formulations used by a well-medicated country like Norway we have 60,000 drugs and formulations with now another 20,000 ayurvedic preparations designed to evade the Drug Controller's scrutiny. Since OPPI, the mouth piece of the foreign multinational pharmaceutical industry, states that the present drug production only reaches 20 per cent of the population, the pharmaceutical industry now demands permission for producing Rs. 16,000 crores of drugs by the end of the century when it is estimated that about Rs. 1,500 crores would satisfy the country's entire need, if used correctly and ethically. That such expansion is being considered favourably by the Ministry of Chemicals which decides the type and quantity of drug production (and not the Health Ministry) shows the extent of the malaise in our health as well as our political system. Despite such overproduction, price-controlled essential drugs for National Disease Control Programmes are always in short supply.

**Table II**  
*Health infrastructure (1981)*

|               | Rural<br>(percentage) | Urban |
|---------------|-----------------------|-------|
| Population    | 70                    | 30    |
| Allopaths     | 27                    | 73    |
| Other system  | 65                    | 35    |
| Nurses        | 31                    | 69    |
| Hospital beds | 13                    | 87    |

While the government allocates Rs. 5,000 crores for public health services equivalent to 1.9% of our GNP, studies by ORG and by FRCH in Jalgaon and Madhya Pradesh reveal that the public is expending at least three times this amount in the private sector, almost exclusively for

curative services. This makes the total expenditure on health care in India to be well over Rs. 20,000 crores which is about 6% of the country's GNP or about Rs. 235 per capita per annum (Table 3). Several studies in our own country like those at Jamkhed, Munnar and our own at Mandwa reveal that far more effective health care can be provided to all our people at far lower cost, for it is not the amount that is expended that matters but the way in which it is spent. While the private sector, especially in cities, is proving to be highly exploitative even the Rs. 60 that is spent per capita by the public sector shows a marked preference for large hospital based curative services in urban areas. Even the meagre Rs. 27 per capita spent on the rural Primary Health Centres (PHCs) is almost entirely devoted to achieving Family Planning targets. This single minded obsession has so alienated the PHC from the people that it has been unable even to achieve its Family Planning goal as revealed by the constant population growth rate of 2.1 during the last three decades.

**Table III**  
*Annual health expenditure*

|             | Total<br>(in crores) | Rural<br>(percentage) | Urban |
|-------------|----------------------|-----------------------|-------|
| Population  | 84                   | 70                    | 30    |
| Expenditure |                      |                       |       |
| Public      | 5,000                | 25                    | 75    |
| Private     | 15,000               | NA                    | NA    |
| Total       | 20,000               |                       |       |

Despite the spawning of a vast manpower, infrastructure, as well as the drug and instrument industry in the past two decades, communicable disease like tuberculosis, leprosy, gastroenteritis, measles, acute respiratory infections, tetanus, rheumatic heart disease, poliomyelitis and filariasis continue to take their unrelenting toll of life and limb; diseases for which we now have remarkably effective knowledge and technology for both control and cure. Tuberculosis continues to affect 9 million persons each year despite a special national programme and claims an annual toll of 400,000 lives. Yet the medical profession and what should now be termed, with the addition of the pharmaceutical and instrumentation input, as the health industry, is observed with esoteric technologies like amniocentesis, *in vitro* fertilization and diseases like cancer, heart and stroke for which relatively little can be achieved at great

cost. Many of these technologies also create economic as well as ethical problems. Preventive and social medicine and family practice have the lowest status in medical education while 60% of the students qualify as superspecialists in increasingly narrow and abstruse subjects. The general surgeon and general physician as of the past is no more produced. There is not a single family physician in our medical colleges when this should be the mainstay of medical education and practice. Like in the West even this is sought to be converted into another speciality while a three-year licentiate course, as before Independence, would be more in conformation with the actual health and medical needs of the majority of our people.

The reversal of priorities from cheap and effective preventive, promotive and basic curative health care for the common disease of the younger age group of the rural poor to expensive specialized curative medicine for the aged urban rich is a perversion of the Bhore model to which the country was committed. This is an example of such departure in every field of our national endeavour from the development model adopted by the father of our nation.

Since neither the public, private nor the NGO sector is able to 'deliver' health care and has only created dependency on this new inappropriate model the question arises as to what may be the alternative. We are informed that there is no alternative but having 'more of the same', i.e. more doctors, more specialists, more urban megahospitals, more drugs and more expenditure. Do they not know that the health status of the USA which spends \$ 800 billion or 12% of its GNP on so-called health care which is equivalent to \$ 2,566 per capita per annum, as measured by the Infant Mortality rate of 9 is only marginally better than that of Kerala which spends less than \$ 15 per capital and whose IMR is 17 (Table IV). The present Structural Adjustment dictated by the same country through its surrogates the World Bank and IMF, states that since the public sector is inefficient we must also privatize it as in the USA. This is a means of converting public assets into private profit which will not only increase the cost of health care, as of all other commodities, but also eliminate the poor from receiving what little they get from the present system. With the simultaneous enforcement of Intellectual Property Rights the price of essential drugs for the common diseases of poverty like malaria, leprosy and tuberculosis will also increase five to ten fold according to our Minister of Chemicals; this just when we expect massive deaths from these diseases as a result of malnutrition due to the inevitable inflation resulting from such perverse antipeople policies.

This grim scenario of the last 20 years is the inevitable result of what has happened to countries following independence who permitted themselves to become enmeshed into the periphery of the capitalist economies. In the presence of the USSR, military adventurism gave place to economic recolonization using the local elite as willing intermediaries in return for an affluent Westernized materialistic life style amidst a sea of abject poverty. The scenario as described in the field of health is a direct result of cultural enslavement and economic recolonization of all those countries after independence who fell into this well designed trap. The economic and health status of countries like China, Cuba and even of Nicaragua are shining examples of what could be achieved by following an independent path based on their countries requirement and not dictated by foreign interventionists (Table IV).

**Table IV**  
*Comparative Data (1991)*

|        | GNP<br>per capita<br>(US \$) | IMR | CBR | Literacy<br>(%) |
|--------|------------------------------|-----|-----|-----------------|
| India  | 340                          | 92  | 30  | 52              |
| Kerala | 340                          | 17  | 19  | 91              |
| China  | 350                          | 29  | 22  | 73              |
| USA    | 20,910                       | 9   | 17  | 96              |

IMR = Infant Mortality Rate

CBR = Crude Birth Rate

Sri Lanka and Kerala have also demonstrated marked improvement in their health status even within a democratic setup. This was the result of a strong political will to serve the interest of the people at large. They not only incurred higher expenditure on the social sector like health and education but also ensured its proper utilization in an egalitarian manner.

Strange as it may seem, in our preoccupation with the problems of the public, private and NGO sectors for 'delivery' of health services we have almost entirely ignored the ability of the people to look after their own health problems. This is partly due to the mystification of health and converting it into an unnecessary complicated illness business. In this process not only the people but also the medical profession has mesmerized itself by the endless import of expensive glamour technology from their mentors in the West. Little do they realize that the highest

technology that Western medicine has to offer is for the prevention and control of communicable diseases. For malaria, tuberculosis and similar diseases we have the precise knowledge of its causation and spread, and a few tablets for ensuring cure. Yet for diseases like cancer, heart and stroke Western science knows little about the cause and even less about its cure. The technology has little to offer except glamour and fear at high cost. This is which I would designate as low technology. And yet it is for these same diseases, Western medicine now prescribes the adoption of a healthier life style as a better alternative. Even this has been converted into a lucrative business such as tranquilizers, food fads and health clubs. Ayurveda and Yoga which have far more to offer at virtually no cost for these sciences knew the adverse effects of unhealthy living, of diet and of mental tension long before Hans Seyle discovered the adverse effects of stress. Yet our medical profession prefers to use drugs, chemicals and complicated surgery which is now being questioned even in the countries of their origin. While this makes good business it is bad medicine. The reason for the reluctance of our medical profession to use our own indigenous systems is also a part of the arrogance of the science and technology which emanates from the West which looks down upon all other systems and cultures. This also affects all those who utilize this science and technology even if it means denigrating one's own culture and practises. Like almost everything that is imported ad hoc it also curbs local initiative and originality and leads to slavish utilization and constant dependency on the West for new ideas. This is the bane of Western science as practised in our country, and is epitomized in the field of medicine. Despite being the largest store-house of disease in the world there has been a remarkable dearth of new discoveries and the constant refrain is for being the 'First in India'. I hope that our younger once released from their mental shackles will look at our own problems with their innate originality so that they may be able to claim to be 'first' without the suffix 'in India'.

Let us for a moment look at the health and illness problems of our country without preconceived ideas and prejudices. It does not require much science to realize that health is essentially a function of nutrition, water supply, sanitation, housing and environment and that most illnesses are the result of lack of these basic human necessities. These are essentially problems that can be solved only by the people themselves through their own political action, though the medical profession can play an important role in bringing this to public attention. The vaccine we

require is not the triple antigen of UNICEF but a multivalent vaccine against poverty.

Even the problems of curative medicine and the use of available technology take an entirely new dimension if it is approached from a practical rather than a merely medicalized point of view, if diseases are classified not according to their pathology but the knowledge, skills and facilities required for their diagnosis, prevention, treatment, cure and control. Experience in health care projects such as our own at Mandwa has demonstrated the advantages of differentiating diseases by utilizing such a practical and functional approach.

Broadly speaking medical problems can be grouped in the following five categories:

#### CATEGORY A

A PHC study by the National Institute of Mental Health and Neurosciences revealed that 25% of all attendance at a PHC comprises of psychosomatic problems which need understanding and explanation rather than a pill or injection.

#### CATEGORY B

This comprises of the commonest simple, self-limiting problems which can be adequately handled by the individual and the family itself, utilizing home and folk remedies and/or some cheap and safe over-the-counter drug like aspirin.

#### CATEGORY C

These are diseases which are not life threatening but are nevertheless responsible for substantial morbidity load of the community, such as scabies, worms, moderately severe diarrhoea, dysentery, acute tracheobronchitis, moderately severe cuts and bruises. They can be adequately diagnosed and looked after by a properly trained Community Health Worker with a modest repertoire of safe but effective drugs. Even more important, by providing advice for management such as oral rehydration therapy. Such advice would be even more appropriate than those of the doctor, because of the physical and cultural proximity of the people, as well as the low cost.



### CATEGORY D

This category comprises of diseases like severe gastroenteritis and dysentery, acute respiratory tract infection, tuberculosis, tetanus, leprosy, malaria, poliomyelitis, measles, pneumonitis, rheumatic heart disease and sexually transmitted diseases. They are the major killers and maimers in the tropics today despite the fact that we have effective knowledge as well as means for their prevention, treatment and control. And yet the knowledge and technology for this is remarkably simple, cheap and effective, as well as being safe. Yet the medical profession has revealed their inability to control these diseases despite monopolizing health care. Experiments such as at Mandwa, Banaswari and Jamkhed have revealed that even these diseases can be most effectively tackled by the people themselves in conjunction with adequately trained paramedical workers. The problem therefore lies not in the inadequacy of medical science and technology or their availability and cost, but in the inability to reach it to the people through the over-professionalized and over-bureaucratized health system in both the public as well as the private sectors. Four decades of experience under the prevailing conditions dictates that this can only be attained through the community's own efforts closely supported by community-based paramedical workers, who in turn, must be appropriately taught and adequately supported by the medical profession and the health services. Withholding this simple and readily available knowledge and technology merely because of the dangerous consequences of these diseases (especially if not diagnosed and treated in the early stages) and a few untoward reactions to drugs, has proved counter productive.

Many of these diseases can be prevented by attending to nutrition, environment, water and sanitation, besides the immunization of the high risk groups within the community. Even diseases like tuberculosis and leprosy can be readily suspected by trained community health workers and paramedics and referred to the doctor for confirmation of diagnosis. The prescribed treatment regimen can then be given and supervised by the local workers, and the patient referred back for occasional checkup or if any of the well known complications of the disease or untoward effects of treatment should arise. In situations where medical help is unavailable, they can be treated by trained community workers with short courses of drugs like chloroquine, enteroquinone, sulphonamides, metronidazole, oral penicillin and can be referred only if the seriousness of the condition warrants. The role of the medical profession must be of a guiding,

encouraging and supportive and not of an appropriate nature as demonstrated by Drs Abhay and Rani BANG in Gadchiroli. Professionals however highly trained have clearly demonstrated their inability to achieve the desired result using the latter approach at vastly greater expense. This is often difficult for the profession to appreciate, in view of the personalized, curative nature of present day medical education and practice.

### CATEGORY E

Comprises of the few high profile problems which need skills and facilities that can only be provided by the medical profession and the hospital. These include major medical and surgical problems and emergencies which are beyond the scope of the para-medical workers who can nevertheless be taught to provide adequate first aid before referral, as well as follow-up and after care. Early diagnoses by such workers, as for cancer, tuberculosis and leprosy, can greatly reduce the load and expense for the treatment of advanced disease so often seen in our hospitals, besides being far more humane as far as the patient is concerned.

It is realized that in a few cases in each category the severity of the disease may increase and upgrade it to the next higher category or categories. A holding period of 48 to 72 hours is generally indicated following which persistence or exaggeration of symptoms indicates referral to the next level of care. The paramedical workers can be taught the signs and symptoms of a few diseases/problems like meningitis and abdominal pain or injury where immediate referral is necessary when in doubt.

Such an approach may seem to smack of over-simplification and even dangerous practice to those who work in countries where quality medical services are readily available to all. Yet a similar attitude by the medical profession in countries where even elementary medical services are inaccessible to most has been extremely counter productive because it denies available medical knowledge and technology to the majority of the people. This approach also leaves relatively few problems requiring greater knowledge, skills and facilities for the attention of the medical profession who can then devote much more of their time as well as the limited expensive resources for secondary and tertiary care for which they are specifically trained. The danger of not utilizing the available knowledge and technology far outweighs the danger of withholding it,

certainly under the prevailing circumstances and even where the socio-economic conditions are better.

The present approach has only succeeded in mystifying health and medical care and placing it within reach only of the affluent minority who can afford to pay for services, or through influence, monopolize whatever effective services are provided in the public sector. The medical professionals, who increasingly perceive medicine as a lucrative business have also mesmerized themselves into believing that 'West is Best'. They have alienated themselves from the masses as well as from their socio-economic problems and culture. The result is that they firmly believe that there is no alternative to the present system, despite its proving to be increasingly counterproductive. The pharmaceutical and medical instrumentation industry, whose sole motive is profit, has spared no effort in reinforcing the belief among the professionals and the public that modern health care must follow the latest Western pattern, however expensive and inappropriate it may prove even in their countries of origin.

### THE COMMUNITY HEALTH CARE SYSTEM

The ICSSR/ICMR report "Health for All — An Alternative Strategy", proposed an alternative model where the vast majority of all preventive, promotive and curative health care would be undertaken by and within the Gram Panchayat and Panchayat Samiti, at the 100,000 population level. This was based on the principle that health lends itself both technically and culturally to be operated most cost effectively in a decentralized system as envisaged in the Panchayati Raj. This is provided that the health and illness approaches are graded and operated according to the skills and facilities, described previously.

Such an approach would involve the people in their own health and help release them from the clutches of the health industry. It would also establish the rapport and confidence between the people and their most important first level contact, the community health workers, without whom no programme can succeed. This will require the provision of more knowledge and availability of essential drugs with these health personnel. This approach can be readily modified to suit the prevailing epidemiologic and social reality which vary from region to region, and often from village to village, in a vast and varied country like India.

The basic unit of such a modified system has to be the Village Health Unit for an average Gram Panchayat of 2,500 population. It will

comprise of five female Village Health Workers and an Auxilliary Nurse Midwife functioning as an integrated unit, supported at the 20,000 population level by a Primary Health Center and a 100 to 200 bedded hospital at the 100,000 population level. The emphasis at all levels will be on preventive and promotive health with the community being involved as a whole. The essence of this system will lie in the community (not merely its elected leaders) having financial and administrative control over the system at all levels from the Gram Panchayat to the Panchayat Samiti. This alone can ensure the desired community participation and involvement as well as accountability without which no public system can function. This is the essence of Panchayati Raj. This together with the hospital with four basic specialities can provide over 95% of all preventive, promotive and curative services of a high level (not merely a second rate service for the rural poor) at about Rs.100 per capita per annum. This is well within the capacity of any government committed to its people's welfare. This has been repeatedly demonstrated by several NGOs in India besides the country-wide experience of China. Also that about 80% of most super speciality functions can be undertaken by broad based specialists with basic medical and surgical facilities at the Community Hospital.

Such a decentralized filtering system with its roots in the community can ensure that about 80% of all problems can be undertaken best at the Gram Panchayat level and about 95% at the Block or Taluka level with intervening Primary Health Centers. The emphasis being on health information, education, prevention and a graded referral service. Such a readily available cheap and effective service would also be the most effective means for the control of the proliferating private sector.

In the intervening decade since the ICSSR/ICMR report was published several projects have reconfirmed the technical, economic and cultural feasibility of this approach; a far superior method to the existing top-down one. The report also clearly indicated that such a health system can function only if the local community is entrusted with its own health and is provided the financial resources and administrative responsibility for operating it. A recent report, (now in the press), termed 'People's Health in People's Hands' provides more detailed information on the Indian experiences for a system eminently suitable for health care in the forthcoming Panchayati Raj.

Such a radical change can only be achieved through the peoples own action. The fears of the medical profession about their own security and dignity are understandable but such a system also offers reasonable monetary as well as job satisfaction to the large number of frustrated doctors and specialists who work in urban hospitals. The job satisfaction that medicine provides with its intense human interplay has somehow been lost in the imitation of the monetized Western model. In such a new order all systems of health care would be utilized with acceptability, effectiveness and cost being the chief considerations.

As a result of disillusionment with the allopathic system due to its high cost, malpractice and exaggerated and unsubstantiated claims, the educated and perceptive middle class had not only taken to suing the medical profession but is increasingly shifting to other systems like Ayurveda, yoga and homeopathy besides self-medication with 'over the counter' drugs. This is also reflected in the sudden increase in the production of ayurvedic drugs, many of dubious quality. This questioning of the allopathic system and a search for alternative systems of medicine is now a world-wide phenomenon. There is also a shift from hospital to home care using family physicians, nurses and paramedicals working within the community. Even among the capitalist countries, except for the USA, health has been protected against the vicissitudes of the market economy and considered a basic human right. Dilution of the National Health Service in the UK by the Thatcher government was resisted not only by the people at large and even the Conservatives but also by the medical profession itself.

It is hoped that the present excesses of the uncontrolled market economy in our country will also lead to a movement towards a more humane and socialized medical and health care, for health cannot be left to the tender mercies of a market economy devoid of moral considerations and in a field where consumer resistance is at its lowest. Such a change cannot occur in isolation in health. It will have to be a part of the overall release of our society from the present grip of mercenary development for the benefit of a select few so that we may revert to a more civilized state in keeping with our culture which places man above money and humility above arrogance.

I have tried to demonstrate, through the field of health, how Western science and technology unless used with the greatest circumspection can lead to severe damage to the health of people under the guise of 'Health

for All'. This science carries with it the ruthless exploitative character of the culture in which it emanates. It has provided those who have originated, and even in those who have borrowed this science and technology, with power to enslave and exploit the rest of the human race and even nature itself to satisfy their insatiable greed devoid of social, ethical and moral restraint. This has prevented the utilization of the enormous beneficial potential of this science and technology which is often dangled as a bait to unsuspecting individuals and societies who are eventually recruited into the process of exploitation of their own people and their own country's natural resources.

This science and technology also demonstrates a new type of fundamentalism which demands that everyone else must believe that it is the repository of all knowledge and technology and all those who question it and its methods and achievements are to be condemned. In this it uses the enormous propaganda machine provided by its mass media to assault and enslave the mind not only of the impressionable youth but also of our senior scientists. This has not only crippled our thinking and originality but makes us look up to and conform with the dictates of this alien society and its science; a society which converts everything into a marketable commodity to be sold to the highest bidder regardless of its consequences to others. And yet Joseph Needham states in his book on *Science and Civilization of China* that "almost half of the principles on which modern Western Science and Technology is based were described hundreds and possibly thousands of years ago in China which is often ignored or denied by the West." This we know is equally true of science in India and certainly in the realm of medicine. It is time we appreciate and utilize our own heritage together with the best of Western science and medicine. This mindless Westernization is why cheap and effective small scale development with a human face has been replaced by expensive large scale dehumanizing business and industry. The ease with which this blatantly false mode has been accepted ad hoc for our country's development by our professional and political leadership, regardless of the entirely different social, economic and cultural needs of our people, is a measure both of their gullibility as well as inability to control personal greed for power and material wealth. This has marginalized the vast majority of our fellow beings in rural India and its urban slums and reduced them to a state of destitution as well as utter dependency and subjugation. They bear the double burden of both international and intranational exploitation.

There are sufficient examples in our own country of what can be achieved by the release and harnessing of our rich human potential. Kerala, Mandwa and Ralegan Shindi are glaring examples of highly cost effective methods for development using local resources and without recourse to international users who then dictate the reverse policies. Our people suffer unnecessarily because their so-called leaders have lost both integrity and nerve. We have tried long enough to solve what are essentially human and social problems through modern techno-managerial fixes. It is time we go to the people so that we can blend our knowledge with their wisdom which is based in our country's reality and use the best of both.

Virchow a renowned German pathologist stated in 1846 that 'medicine is social science and politics is medicine on a large scale.

Finally let me quote our own scientist C V Raman who said that "The quality of the Indian mind is equal to the quality of any Teutonic, Nordic or Anglo-Saxon mind we have. I think we have developed an inferiority complex. So what is needed in India is the destruction of that defeatist spirit."



**Balasubramaniam Ramamurthi** (b. 30 January 1922) Madras, Tamil Nadu M.S.(1947) from University of Madras, F.R.C.S.(1947), Edinburgh, UK He is Head, Department of Neurosurgery, Voluntary Health Services Medical Centre, Madras; and formerly Professor of Neurosurgery, Government General Hospital, Madras (1958-78); Principal, Madras Medical College, Madras (1976-78); Visiting Professor at eight foreign and twelve Indian universities

Ramamurthi's researches span injuries to nervous system, surgery of brain tumours, tuberculous infection of brain, neurophysiology, stereotactic surgery, consciousness and biofeedback techniques. A study of neurophysiology of the deep structures of the brain he made during stereotactic surgery led to a better understanding of movement disorders, pain and epilepsy Investigated the yogic states of meditation and established the value of biofeedback techniques Also studied the correlation of neurology with ayurveda and yoga Developed techniques of prevention and treatment of head injuries

Ramamurthi is Fellow of International College of Surgeons, American College of Surgeons, Indian Academy of Sciences, and National Academy of Sciences (India); President, National Academy of Medical Sciences (India), Neurological Society of India, Association of surgeons of India, Neurosciences Society of India, Indian Association of History of Medicine, Indian Association of Biomedical Sciences; Hony Member, Australian Society of Neurosurgeons, Society of Neurological Surgeons (USA), American Academy of Neurosurgeons, Japanese Neurosurgical Society, Scandinavian Neurosurgical Society and Congress of Neurological Surgeons; Hony President, Asian and Australasian Congress of Neurosurgery and World Federation of Neurosurgical Societies; Member, INSA Council (1988-90). He is the recipient of INSA-TS Tirumurti Memorial Lecture Award (1985), Hony Surgeon to President of India, Dhanwantari Award Hony Brigadier, Padma Shri, Padma Bhushan

*Balasubramaniam Ramamurthi was elected to the fellowship of the Academy in 1981.*

**Address:** Head Department of Neurosurgery, Dr. Achanta Lakshmipathi Neurosurgical Centre, Voluntary Health Services, Adyar, Madras 600 133; Tel: (044) 2352018, 2352123; Res: 27 Second Main Road, CIT Colony, Madras 600 004; Tel: (044) 71475/71356.



## **DOES CHEMISTRY EXPLAIN THE MIND?**

B RAMAMURTHI FNA

### **INTRODUCTION**

Mind has been known to man for more than 10000 years and over the past 5000 years mankind has interested itself in determining the functions and the components of the mind. An enormous amount of knowledge had accumulated over the centuries about the powers and the functions of the mind specially in countries like India and China with long-standing civilisations. In India, our ancient rishis and seers have looked deep into their own mind and mental functioning and had enunciated many methods of mind and body control. Thus while our familiarity with the mind is age-old, our knowledge about the human brain is comparatively very recent. Awareness about the functions of the brain and the important and vital role it plays in body and mental functions was realised only over the past 200 years.

### **‘MIND IS MATERIAL’**

In the last century and in the early half of the 20th century, it was often the fashion to discuss the relationship of mind and the brain and mind and matter. The Indian Philosophers and Psychologists have defined mind as a part of our material existence. Mind as "Manas" is defined as one of the Indriyas (Geetha: Indirayanam Manaschasmi). In Indian Psychology mind is not a single entity but consists of different layers of hierarchy namely the mind or the manas, buddhi or intelligence and ahankara (self awareness). From this it is clear that the Indian psychologists were way ahead of their western counterparts in their concept of the mind and in their understanding of the role of the different aspects of the totality known as the mind. The Indian psychologists also surmised that these mental layers could expand beyond the confines of the individual mind and embrace the "universal mind" but that is another story. There were many philosophers and psychologists in the West who held that the mind was something entirely different and that it had a separate immaterial existence different from the brain, while others tried to correlate the concept 'mind' with the organ "the brain".

## MIND IS A CONCEPT

The term 'concept' has to be used regarding the mind as this entity is only a concept in the thinking of men. All organs in the body have functions and words have been used to identify such functions e.g. digestion is the function of the alimentary system; breathing is the function of the lung; movement is the function of the muscle and so on. Similarly the mind or mental functioning is one aspect of the function of the human brain.

The dependence of the mind on the brain is easily clarified. It is obvious that whenever the brain is affected in any manner, the mental function immediately changes. In instances where the brain is seriously affected or damaged the mind ceases to function e.g. after a severe head injury the patient loses consciousness and the mental function ceases. He is not aware of his environment nor is he able to respond to the same. Similarly after the administration or intake of certain drugs the patient may lose consciousness as happens when an anaesthetic drug is given. Changes in the mental functioning like memory intelligence and judgement occur when the oxygen and blood supply to the brain are impaired as may happen in diseases or in old age. Similarly when tumours occur in certain areas of the brain, the patient loses his ability to think coherently, he may lose his memory and in the later stages may become unconscious. It is obvious therefore that proper functioning of the brain is most essential for mental status.

It will be the purpose of this oration to try and define the relationship of the brain and its function namely the mind and to try to present modern knowledge about the electro chemical activities in the brain which determine the mental status and mental powers of the individual.

## NEUROSCIENCES RESEARCH

During the past 200 years there has been an enormous expansion of our knowledge about the functioning of the nervous system and the human brain. The advances have been specially rapid through the past 2 or 3 decades during which time neurosciences have attracted many brilliant intellects. These include not only neurologists, neuro-physiologists and neurochemists, but also psychologists, social scientists, computer engineers, electronic experts and mathematicians. Such concentration of effort has yielded a vast amount of data regarding the functioning of the human brain. Still it has to be accepted that what is known is very little

and there are many problems to be solved even about the normal functioning of the human brain.

### THE MARVEL OF THE HUMAN BRAIN

The human brain, the most marvellous creation, consists of more than 1000 million cells which are termed neurons in addition to having a similar number of supporting cells called glial cells. Each neuron has a long process the axon along which it sends messages and a number of small projections called dendrons through which each neuron receives messages from other cells. There could be anything between 10 to 50 such connections between different cells. One can imagine the number of permutations and combinations that are possible for energy transfer through a system consisting of 1000 million units or more and where each unit can make anything between 10 to 150 connections. The possibilities are really astronomical.

There is an intricate and constant and instantaneous consultation taking place between various parts of the brain and the periphery, between the right and left brain through cross connections, between the frontal, the temporal, the parietal and the occipital portions of the brain on the same side with each other and with their counterparts on the opposite side and with lower concerned levels in the organisation. All this interchange of information occurs within milliseconds and the right answer to help to preserve the integrity of the organism presented in less than a second. This conglomeration of enormous numbers of nerve cells situated inside our head and termed the brain controls every function in the human body by fibre connections that establish contacts with the skin and with all organs, muscles and all other structures. There is no function performed by the organism from the highest thinking to the basic functions like breathing and heart beat which is not controlled by the brain.

The brain has come upto this present state of complexity over evolution through millions of years by building up from simplest mechanisms and by overlaying of more complex functions on simpler functions, thus building a hierarchy of functional integration. You may be surprised to know that the lower forms of life on earth have perfected many intricate neural and communication responses, like sonar in bats, communication between dolphins, photosignals, mating signals etc.

## COMPLEX ELECTROCHEMICAL CONGLOMERATE

The functions of the central nervous system are mediated by chemical processes. Every organism needs for its existence the ability to obtain and to react to information from its environment and from within itself. In the unicellular organisms this is achieved by simple chemical processes, example chemotaxis, an organism is attracted towards its food or repelled from an unsuitable environment. In multi-cellular organisms it was not only important to communicate within the cell and with the environment but information processing became necessary between the cells themselves. As the complexity of the organism became greater, certain cells began to specialise in information transmission and collection, evolved into the highly complicated neurons and established contacts with all the cells in the body and with the periphery. With the increasing complexity of the nervous system resulting from the development of varied and complicated functions in the higher animals and the human beings, it was necessary to have different systems of information transmission and processing; rapid transmission for quick results, slower action for diffused results and much slower action spread over the years. This was achieved by different types and speeds of transmission of neural impulses by adopting different electrochemical mechanisms.

As far as our present knowledge goes, energy generation in the brain is through an electrochemical complex. Shifting of the sodium and potassium ions from the inside and outside of the cells creates an electrical potential which travels down the axons and initiates action in another neuron or in an organ like muscle or gland.

The impulses generated by the neurons influences other neurons through what are known as synaptic junctions. These synapses are junctions between the nerve fibres and the nerve cells through which the electrical impulse coming along with nerve fibres get transmitted. *The synapse is the most important structure in the nervous system* that determines the functioning of the brain.

## CHEMICAL MESSENGERS

The modus operandi of this complicated system has been investigated for many decades and recently rapid progress has been made with the discovery of new chemical messengers and mechanisms. These messengers have been termed neurotransmitters, neuromodulators and neurohormones. They are of differing chemical compositions.

Our knowledge of the functioning of neurotransmitters has enlarged because of important technical advances in the methods of selective staining of neurons containing the particular transmitter. These techniques are based on fluorescence or radioactive labelling or on specific antibody techniques and have provided information about the areas of concentration of the individual transmitters in the complex neuronal circuits of the brain. It is found that the transmitters are not distributed diffusely but are precisely localised in specific centres and pathways.

## NEUROTRANSMITTERS

One of the earliest neurotransmitter to be discovered as playing an important part in neuronal function was acetylcholine. This chemical is present in the brain and the central nervous system and also in the periphery at the junction of the nerves and muscles (neuromuscular junction) where it plays a vital part in inducing muscle contraction. Interference with the acetylcholine mechanisms in the periphery cause a serious disease known as myasthenia gravis. Acetylcholine deficiency in brain may cause premature loss of mental function (presenile dementia).

The other transmitters involved in brain and nerve functioning are noradrenaline, serotonin and dopamine. The noradrenaline system has a reciprocal relationship with the acetylcholine system with opposite functions. When one system acts like an accelerator, the other system acts like a brake. The noradrenaline system is necessary for defence mechanisms in the body. When an animal is threatened, it has to defend itself by fighting or running away (fight or flight). For this purpose, large amounts of energy stores have to be released urgently and the noradrenaline system is most essential for this, to get more oxygen and blood to the tissues, to make the heart pump blood faster, to provide the necessary glucose for muscles to act etc. The hypothalamus in the brain is the chief main centre for this functional mobilisation.

In ordinary circumstances the noradrenaline is utilised by the fight or flight mechanism. But in modern society, problems are not solved by a simple solution of fight or flight. In many situations when anger arises one is forced to remain silent. This leads to frustration and the unused noradrenaline leads to psychosomatic illnesses like peptic ulcer, blood pressure, headaches, neck and back pain insomnia etc. These are the diseases of fashion and frustration. The advice 'not to get angry at all' is good. If one gets angry it is better to do some physical activity to vent out

the anger rather than suppress it. No wonder husbands went on long walks when there were domestic quarrels. How about the wives? They made pots and pans fly.

The dopamine and serotonin systems are inside the central nervous system. The dopamine systems control mental process as well as muscle movements. In schizophrenia, the dopamine system is disturbed. Serotonin systems are involved in functions like alertness, sleep, temperature regulations etc. Serotonin excess or deficiency may cause certain problems and one of the well known problems is migranious headache.

### INHIBITION OR SUPPRESSION IN THE BRAIN

It was for many years assumed that the nerve cells in the brain show only a positive response i.e. a stimulus results in a discernible or a measurable reaction. However it was found that there were instances where such a reaction was absent and also some instances where a stimulus resulted in an effect being slowed down or stopped. This led to further investigations of the functioning of these neurons and it was discovered that there were many cells in the nervous system that inhibit or slow down reactions rather than start or speed up a process. Such inhibitory cells are known to be present throughout the nervous system and the important role they play in protecting the nervous system and in making it function optimally has been fully recognised.

Such inhibitory groups of cells act as 'gates' that allow or do not allow impulses to pass through. This should have been expected, as otherwise, millions of stimuli bombard the human brain every moment and it will be virtually impossible for the brain to function without such protective devices. We can see many instances in every day life where such protective or inhibitory mechanisms are used e.g. one can work and concentrate in the middle of disturbances. If one is sufficiently interested, one can study or write in spite of a distracting loudspeaker noise on the street—a common phenomenon in our country.

How is the brain protected from such disturbances and allowed to work? The inhibitory cells are placed from the most peripheral parts of the nervous system to the very centre. Inhibitory or gate mechanisms are seen in the skin and organ surfaces, in the peripheral nerves, in the ganglia near the spinal cord, in the spinal cord itself, in the lower levels of the brain known as the brainstem, in the cerebellum and also distributed throughout the cerebral cortex. These cells make it possible for the brain to function in

spite of bombardment by innumerable stimuli. The inhibitory transmitter involved is gamma amino butyric acid (GABA). This inhibitory transmitter is manufactured exclusively in the brain and the spinal cord. GABA mechanisms are involved in most of the activities of the brain. More than a third of the synapses in the brain employ GABA as a transmitter. Recently GABA has been implicated as a probable target for the actions of drugs like Diazepam.

### NEUROHORMONES—SLOW ACTORS

The chemical mechanisms that we have talked about are used by the nervous system to produce immediate and instantaneous action in the periphery, the time scale being in millionseconds. The brain also has to function at a slower speed in certain situations and also at a very slow speed spreading over many months and years in certain areas: e.g. in sexual maturation and growth of the organism, the brain plays its part by very slow acting mechanisms and achieves its aims through glands situated in various parts of the body known as the endocrine glands and the chemical messengers are neurohormones. The brain controls the hypothalamus which in turn controls the pituitary gland, the adrenals, the ovaries and the testes. Such neurohormonal systems also control the time clock mechanism situated in the hypothalamus that determines when a child becomes a sexually matured person and also determines the growth in height. Of course there are also short and swift acting endocrine mechanisms in this system which are essential for sexual function and also for the growth and preservation of the life of the fetus inside the mother.

### SYNAPTIC TRANSMISSION

Precise biochemical events are involved in synaptic transmission and when such events are interfered with either by injury, disease process or toxic substances, the mechanisms in the brain get altered resulting in drowsiness, excitement psychosis etc. Drugs are now available that selectively enhance, inhibit or block specific steps of synaptic transmission. This has given us good knowledge about the mechanisms of action of drugs used in mental diseases and also an idea of how some neurological and mental disorders might be related to specific defects in the mechanisms of synaptic transmission

A number of steps are involved in the process of chemical synaptic transmission. Synthesis, storage, release, reaction with the receptor, termination of action and reabsorption or destruction.

The molecules of the transmitters are manufactured by a series of enzyme reactions in the body of the nerve cell from the constituents circulating in the blood stream. At the end of the axon (long nerve projection) they are stored in small quantities in tiny sacs called synaptic vesicles. When the nerve impulse arrives at the axon terminal, large numbers of transmitter molecule are discharged into the synaptic space and travel rapidly across this space to interact with specific receptor sites on the other side of the synapse called the post-synaptic membrane. Once the interaction has taken place, the excess of transmitter molecules are destroyed by an enzyme or reabsorbed into the axon terminal.

The receptors for the neurotransmitters are large protein molecules embedded in the cell membrane with parts sticking out like projections. These projections are precisely tailored to match the configuration of the neurotransmitter molecule so that the two molecules fit into each other with precision. This interaction may cause a neuron to become excited or inhibited, a muscle cell to contract or a gland cell to secrete a hormone. Some reactions take place within a fraction of a second while others like secretion of hormones may take minutes or hours.

### PSYCHOACTIVE DRUGS (Drugs Used in Mental Illness)

Modern psychoactive drugs act by interfering with the mechanisms of synaptic transmission. They may enhance or inhibit the release of a particular neurotransmitter. Amphetamine (Dexedrine) increase the release of dopamine in the nerve terminals in the brain and enhances arousal and a feeling of well being. Excess of amphetamine leads to hallucinations and delusions, resulting probably from a derangement of dopamine systems. Drugs like chlorpromazine and haloperidol bind tightly to dopamine receptors in the brain, thus preventing the action of natural transmitters on the receptors. These two observations on the mechanisms of action of amphetamine and chlorpromazine has led to the suggestion that schizophrenia could be due to disturbance of dopamine systems.

Other psychoactive drugs may act by mimicking the natural transmitters. The hallucinogenic drug mescaline is similar to dopamine and norepinephrine, and LSD is related to serotonin. Drugs like caffeine and theophylline act by increasing the amount of cyclic AMP that is produced in response to a stimulus, thus exerting a general mild stimulant action on brain. Other drugs inhibit the enzyme monoamine oxidase (MAO) which degrades norepinephrine. Thus MAO inhibitors act as



antidepressants. Tricyclic antidepressants like imipramine and amitriptyline act by blocking the reuptake of NE and serotonin from the synapse. Cocaine also acts by a similar mechanism. It is thus clear that by manipulating the synaptic transmission mechanisms in the brain, *it is possible to alter mood, behaviour, thought processes, sleep, attention etc;*

## NEUROPEPTIDES

In recent years our knowledge of the number of chemical messenger systems known to function in the brain has expanded dramatically with the discovery of neuro-peptides. These are short chains of amino acids that are widely but precisely distributed in the central peripheral and the autonomic nervous systems. The occurrence of neuropeptides in specific cell groups strongly suggests that they influence and control specific neural functions. Approximately 50 neuropeptides have been identified in invertebrates and are found to control behaviour in these organisms. There is no doubt that neuropeptides form a basic mode of communication and control in the nervous system. Neuropeptides found active in the highly evolved human brain are also seen in the lowest forms of life, *thus emphasising the fundamental nature of the biochemical processes involved in neural control.*

Neuropeptides may function as neurotransmitters in the classical sense like acetylcholine and dopamine. Others may act to modulate the action of another neurotransmitter. They also produce diffuse changes in many groups of neurones, thus influencing complex behavioural patterns. The chemical sequence of amino acids in each neuropeptide has been identified and by the use of immuno cytochemical and radio immuno assays their site of production and action have been determined. Some of these neuropeptides originally found outside the nervous system have recently been identified as occurring also in the brain.

The first group of peptides to be discovered were the substances that act on the smooth muscles of the blood vessels, or the viscera; the second group are the hypothalamic releasing peptides and the newest group are the encephalines and endorphins.

## SUBSTANCE P

Substance P is intimately concerned with pain transmission and perception and is found concentrated in the sensory neurones in the periphery and in certain specific nuclei in the brain and the spinal cord. Though the

existence of substance P was first described in the extract of brain and intestine as early as 1931, its chemical composition and precise action could be determined only when better bioassay systems became available. SP has been implicated in the sustained pain which persists long after the application of a painful stimulus.

### ANGIOTENSIN

Angiotensin is widely distributed in the brain. This neuropeptide has a striking effect on drinking behaviour and seems to be involved directly and indirectly with the maintenance of extracellular fluid volume and the electrolyte balance.

### HYPOTHALAMIC RELEASING FACTORS (Hypophysiotrophic neuropeptides)

These substances originally described as releasing factors for pituitary hormones are now known to belong to the group of neuropeptides. The amino acids sequence for some of these neuropeptides has now been determined. They are (1) Thyrotrophine releasing hormone (TRH) (2) Luteinizing hormone-releasing hormone (LHRH) (releasing both the leutinizing and follicle stimulating hormone) and (3) somatostatin (SOM) which suppresses the release of somatotrophin (the pituitary growth hormone).

The surprising fact about TRH LHRH and SOM is that they are found all over the nervous system and not confined to the hypothalamic regions. Apart from its action through the pituitary, LHRH has a specific activating effect on sexual behaviour which is mediated via the nervous system.

Somastatin has a depressant action on the central nervous system; it enhances barbiturate anaesthesia and induces hypothermia (lowered temperature) SOM has also been found in the gastro intestinal tract and the pancreas.

### ENCEPHALINS AND ENDORPHINS (Opioid Peptides)

The most exciting and the newest of the neuropeptides are the encephalins and endorphins—chemicals occurring naturally in the brain that bear a surprising similarity to morphine. The discovery of these neuropeptides stemmed from the observations that certain regions of the brain were found to bind opiate drugs with high affinity. The opiate receptors are

found concentrated in the areas that are involved in the perception and integration of *pain and emotional experience*.

### ENCEPHALINS

In 1975, two naturally occurring peptides were discovered that bound tightly to the opiate receptors in the brain and were named enkephalins. Both the enkephalins are chains of five aminoacids. The terminal aminoacid is methionine in one peptide and leucine in the other—hence they are termed Met. Enkephalin and Leuencekephalin. Both mimic morphine and related opiates by binding to the same highly specific receptor sites. Most of these receptors are related to pain while some are not. It was soon realised that enkephalins must have many other actions unrelated to pain as they were found concentrated in areas not concerned with pain mechanism.

Beta lipoprotein, a large peptide found in the pituitary contains metencekephalin. Some of the large frangements of beta lipoprotein from the pituitary gland are found to interact specifically with opiate receptors and have been termed endorphins.

Enkephalins are also found in the dorsal horn of the spinal cord at the same sites as substance P. If substance P is an activator of pain impulses, enkephalins are suppressors of pain at the same sites.

### PROSTAGLANDINS

Amongst other chemicals that appear to play a modulatory role in the brain, the prostaglandins are found widely distributed and are present at high levels in the brain tissue. They elicit a variety of excitatory and inhibitory effects on the neurons, depending on the precise molecular structure of the prostaglandin and the nature of the target cell. The prostaglandins effects long term shifts in membrane polarisation and modulate the effect of neurotransmitters.

### GLOBAL CHEMICAL CODING

The neuropeptides are chemical messengers of a nature different from that of the previously identified transmitters. They appear to represent a global means of chemical coding for patterns of brain activity associated with particular functions like sexual behaviour, pain or pleasure, fluid balance etc. Biologically active peptides originally found in the G.I. tract such as gastrin, substance P vaso active intestinal polypeptide (VIP) and

cholecystokinin are also present in the central nervous system. Some peptides originally found in the brain have now been found to occur in the intestinal tract e.g. enkephalins, somatostatin, neurotensin. The fact that many of the neuropeptides are found outside the central nervous system (e.g. the gastro intestinal tract) shows that neuropeptides serve a multiplicity of roles acting as local hormones in the gastro intestinal system and as *global transmitters in the brain*.

For a long time it was not possible to correlate some of the slow processes that occur in the central nervous system, with the rapidity of the electrical process of nerve conduction. With the discovery of the neuropeptides and the elucidation of their mechanisms, specially in the field of neuro endocrinology and behaviour responses, this is better understood.

Vigorous research is going on all over the world in the field of mental functioning, mental life and in psychiatric disorders, to understand the physical and chemical basis and some amount of success has been achieved. In our own country and elsewhere, chemical changes occurring in the brain and in the cerebrospinal fluid have been demonstrated during mind control, mental concentration and yogic states. Chemical changes have also been shown to occur during biofeedback techniques and operant conditioning. In other words there are many nonchemical processes and techniques which can alter the chemical basis of mental functioning and thus influence the mental status of an individual, yoga and other Eastern techniques of mind control being the best of them. It is also conjunctured that the electrochemical potential of the human brain is so vast and enormous that the brain is capable of much more effort and achievement than has been obvious so far during human development.

### CONCLUSION

We have so far seen that the brain is a conglomeration of millions of neurons that functions and controls itself and the entire human organism through its electro chemical activity. This function is mediated by numerous chemical messengers acting in various precise and specific locations in the brain. It appears as if some of these chemical messengers act in one specific area, others have a more diffuse function while certain others assume a global role.

In the time scale of action, some messengers are responsible for swift action, some for slow action and others for very slow long lasting

effects. It has also been indicted that throughout the brain, there are mechanisms and messengers which act as brakes or inhibitors of neural function. It is this marvellous organisation with millions of neurons throbbing with activity, with controls and feed back at every stage and every step, mediated by complicated electrochemical mechanisms, that determines all the known functions of the brain.

While the mechanisms of gross lower functions like self preservation, feeding, sex and reproduction are clearly known, the complex mechanisms responsible for the higher levels of response as in memory, behaviour and intelligence are only vaguely understood. Of course it has been clearly indicted in the preceding paragraphs how chemicals influence varieties of brain functions like consciousness, attention, mood, drive etc. and also the precise mechanisms of such action. But in many areas of higher human development like intelligence, altruism, unselfish love, generosity, kindness, sympathy, sacrifice for others etc, we have no concept at all of how the brain functions either physically, chemically, electrically or otherwise.

Are these qualities inbuilt in the brain and genetically determined. Are the chemical processes in brains different between a criminal a cruel sadist or a kind person.

From the knowledge so far available, can we correlate changes in cell functioning with changes in our mental function? Today the answer is clearly a 'no' as our knowledge is very incomplete. From neurons to brain, from brain to mind and mind to intellect still seems a long way, but the human mind in its enthusiastic quest will enlarge in the near future our knowledge about its own functioning.

## References

1. Iversen L L 1979 The Chemistry of the brain; *Scientific American* **241** 134-149
2. Pullan P T, Clement Lones V et al. 1980 Ectopic production of methionine enkephalin and beta endorphine; *Brit. Med. Journal* **280** 758-759
3. Ramamurthi B 1981 Yoga: Neurophysiological Basis; *Vedanta Kesari* **68** 485-488
4. Ramamurthi B and Vasudevan M C 1982 Yoga and Brain Mechanisms with special reference to inhibition; *The Yoga Review* Vol. II, 61-67
5. Sweet W H 1977 Neuropeptides; *Neurosurgery* **1** 311-316  
— and Smith B H 1978 Neuropeptides and behaviour; *Neurosurgery* **2** 62-69



**Srish Chandra Seal** (b. 4 July 1903) did D.Sc. (1980) from University of Bombay. He is Chairman, Indian Institute of Family Welfare and Health Education and was formerly Emeritus Medical Scientist, Indian Council of Medical Research (1970-73), Professor of Epidemiology, All India Institute of Hygiene and Public Health, Calcutta

Seal's greatest success is on plague, which has been eradicated in India. Another important contribution of his is on cholera by discovering cholera extoxin. A vaccine prepared from this, cholera toxo-vaccine, could help eradicate this disease. He also did much work on epidemic dropsy, which has been conquered. Seal also developed a 'Rural Health Service Scheme', in which all primary health centres have been brought under WHO's 'Health for all by 2000 A.D'.

Seal is Fellow of Asiatic Society, National Academy of Sciences (India), Indian Public Health Association, Indian National Institute of Ecology, American Public Health Association (Hony), American College of Physicians (Hony), President, West Bengal Voluntary Health Association, Rural Welfare Society (Diamond Harbour), and Indian Public Health Association; Foundation Fellow, International Medical Science Academy (Chairman, Calcutta Chapter) He is the recipient of Coates Medal (University of Calcutta) (1952); President, Medical and Veterinary Section, Indian Science Congress (1963), Leadership Recognition in Public Health (American Public Health Association) (1981), Dr T S. Tirumurti Memorial Lecture Award (INSA) (1987)

*Srish Chandra Seal was elected to the fellowship of the Academy in 1955*

**Address:** Research Advisor, Calcutta Family Welfare Hospital Calcutta, Tel. (033) 274787; Res. 2 Fern Place, Calcutta, Tel. (033) 4405541

## HISTORY OF RESEARCHES ON PLAGUE LEADING TO ITS ERADICATION IN INDIA

S C SEAL FNA

### THE PROBLEM OF PLAGUE—HISTORICAL ASPECT

From the evolutionary point of view diseases that were prevalent in the animal kingdom gradually involved human beings by legacy or contact as they started to live in groups or communities. Plague is one of them which even now gets from the animals (rodents) From the available historical accounts it appears that plague visited all regions and countries in the world at one time or the other since the middle ages, leaving a few islands and isolated areas free during the course of known historical period.

The first human epidemic on record according to Wu-Lien-Tech et al. (1936) was the outbreak among the Philistines in 1320 B.C. It was characterised by the appearance of emerods (buboes) in their secret parts, as described in I Samuel V and Vi in the Bible. The Indian scripture like the *Bhagabata Purana* (1500-600 B C) gives the disease an earlier antiquity by referring to deaths caused by an epidemic disease preceded by an epizootic among rats. Men were warned to quit their houses when the rat fell from the roof (apparently *R. rattus*), jumped about as if it was drunk and died. As the time passed the accounts of plague became more reliable. It was prevalent in Egypt, Libya and Syria as found in the writings of Rufus, the physician at Ephesus round about 100 AD.

Of the classical pandemics the *First* one occurred during the regime of the Emperor Justinian (AD 542). It started from Pelusium, the great trading centre of lower Egypt from where it spread through N Africa to Roman empire on the one side and to Syria, Palestine and Constantinople on the other and then to other parts of Europe and Asia reaching London in AD 662. It lasted for 50-60 years and killed 100 million people.

The *Second* pandemic started in the 14th century (AD 1347) from Caffé in Crimea and spread to China and India on the one side and Asia Minor and N Africa on the other. It was imported to Geneva through the Army and from there it appeared in other parts of Europe reaching England by AD 1349. The disease then ravaged Europe for well-nigh three centuries under the horror of what was known as "Black Deaths" taking a

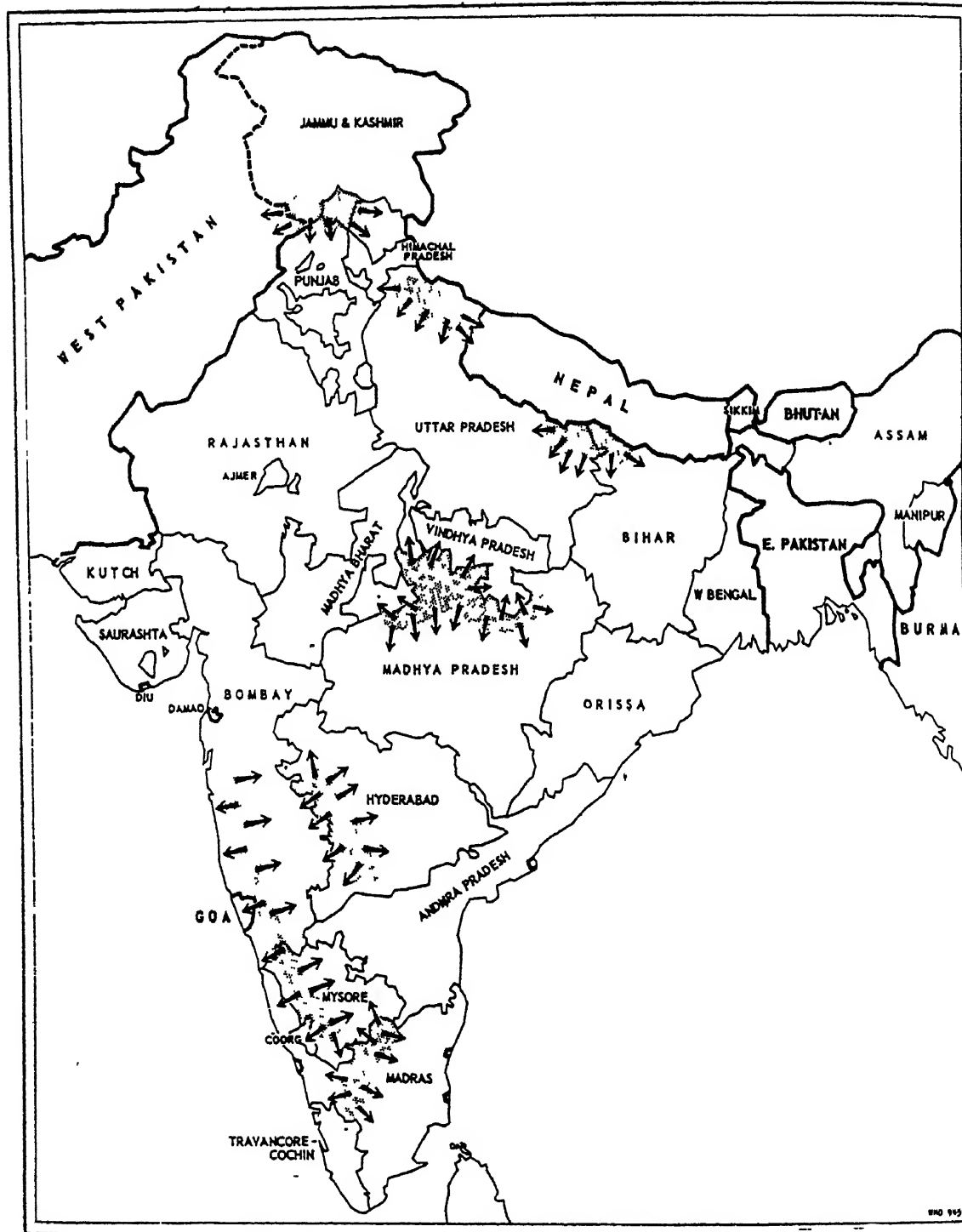


FIG 1 Endemic plague foci in India according to Sharif (1951). Arrows indicate only the direction of radiation of plague, not the actual course of its progress.



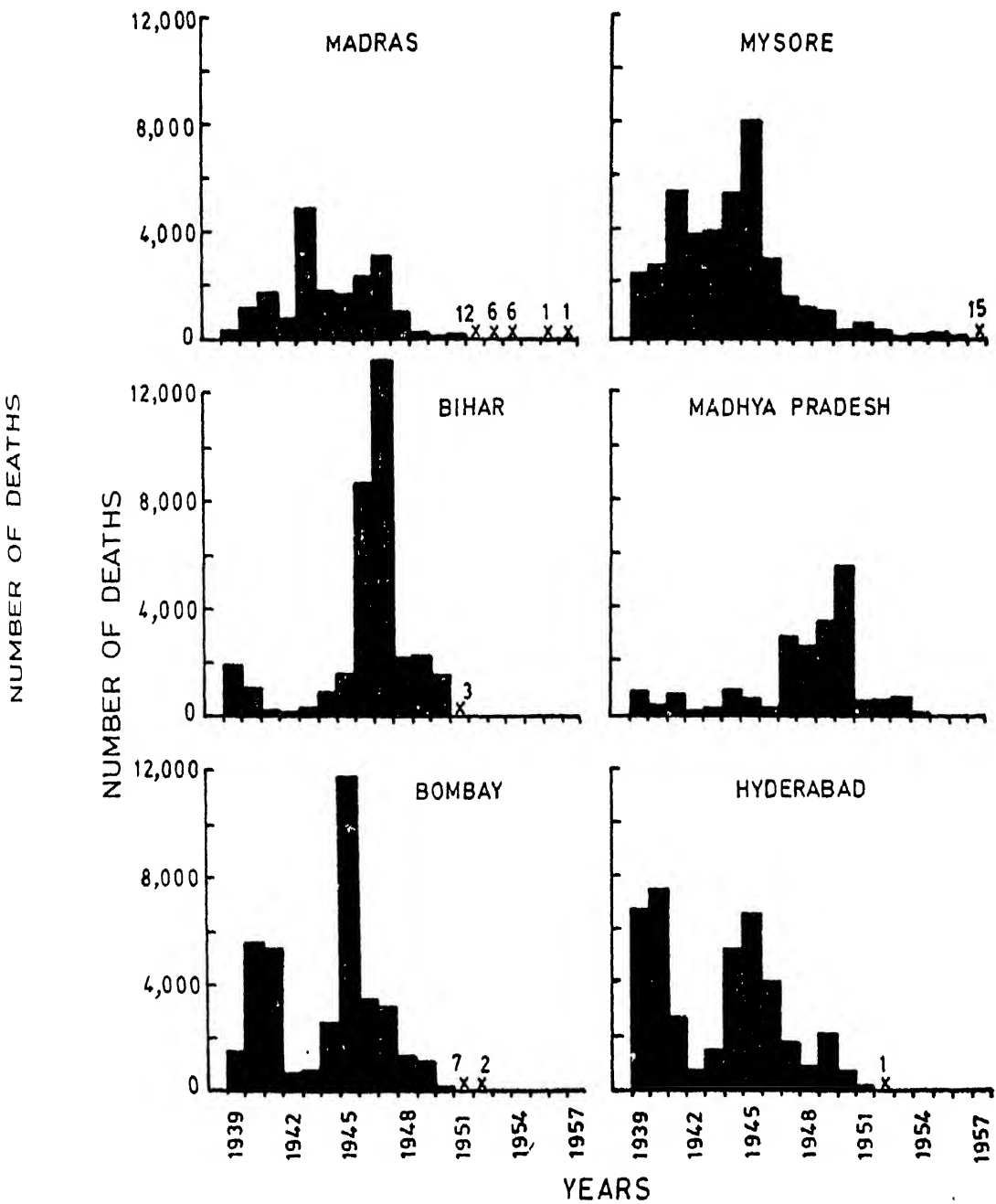
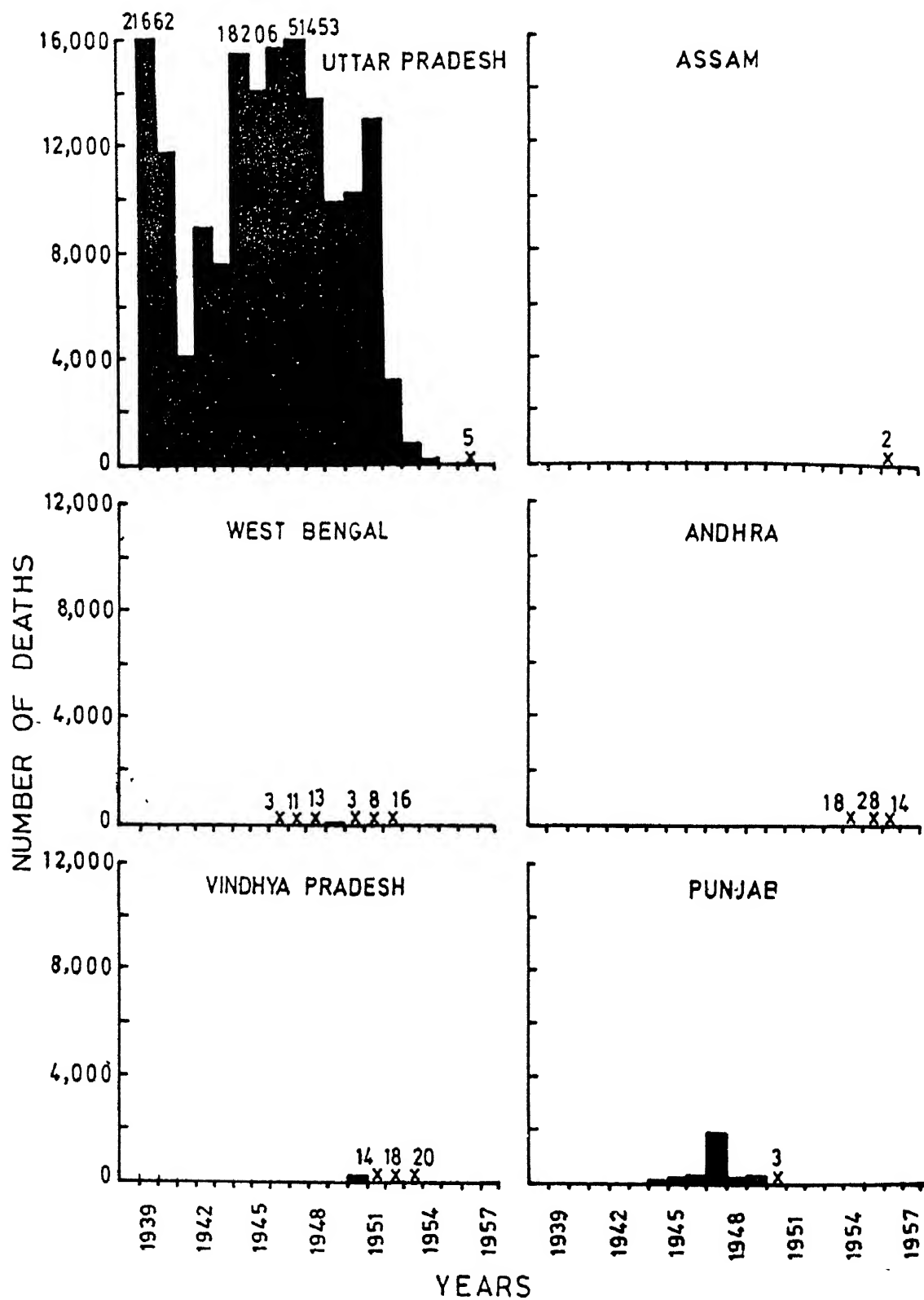


FIG 2a-b Plague deaths in different States in India (1939-57)



toll of 25 million human lives. Retrogression started in the 17th century from west to east completely leaving the European stronghold by 1841. During this time quarantine laws were first passed in 1374 by Count Barnardo of Reggio and by the Venetians in 1403, followed by land cordons and sanitary improvements in the town on the basis of the ideas that rats were involved and the spread was through the maritime transport. About this time Gregory of Tours (540-94 AD) convincingly traced the transportation of plague infection from Spain to Marseilles harbour in 588 AD in support of the view that the Justinian and Byzantine plague in 542 was brought by the corn-ship from Palestine.

The history of plague in India can be divided into two distinct phases namely, Phase I (before the discovery of plague bacillus) and Phase II (after the discovery of plague bacillus). During the first phase reference to plague is available in *Bhagabata Purana* (1500-600 BC) and next to it the occurrence of plague in Malabar in 1325 AD, in Malwa in 1403 AD and in Punjab in 1617 AD during the reign of the Moghul Emperor Jahangir. It was repeated in Kathiawar and Kutch in Gujarat in 1812-21 AD and followed by 'Pali' plague in Rajasthan (1836-1838 AD)

The *third* pandemic started in 1894 AD and was traced to the reappearance of plague in South China at Yun-nan-fu in 1866 whence it reached Hong Kong in 1894 AD and from there to Calcutta (India) in 1895 and Bombay in 1896 by the maritime route. Plague bacillus which was discovered by Yersin and Kitasato in 1894 was isolated on the 17th April in Bombay and 10th October at Calcutta in 1897. From these two cities the disease spread to almost all parts of the West and North but not towards Assam and East Bengal. The author during the course of his research discovered the reason why the disease did not invade these two regions. This epidemic situation prevailed till 1918 with a peak in 1907 and a total death record of 10.25 million. In 1907 itself there were 1,315,892 deaths. In 1910, a great epidemic of Pneumonic plague occurred but by 1919 all plague epidemics began to decline everywhere except Java and East Indies leaving certain persistent foci in India defined by Sharif (1951) (figure 1). It started with a decennial mortality rate of 183.3 per 100,000 in 1898-1908 to 11.7 in 1929-38 and 0.02 in 1959-68 before its fang could be completely broken. Its ingress into different states in India (Seal 1960b) is shown in figure 2, and table 1.

**Table 1**  
*The province-wise distribution of plague mortality between  
 July 1898 and June 1932*

| Province     | Mean population<br>Census 1901,<br>1921 and 1931 | Total plague deaths<br>(July 1898–June 1932) | % of All<br>India total | Mortality<br>per 1000 of<br>mean population |
|--------------|--|--|-------------------------|---|
| Punjab       | 21,142,793                                       | 3,489,123                                    | 28.7                    | 165.0                                       |
| Bombay       | 19,877,756                                       | 2,460,132                                    | 20.2                    | 123.8                                       |
| U.P.         | 47,164,594                                       | 2,911,837                                    | 23.9                    | 61.7  |
| Bihar+Orissa | 34,692,676                                       | 1,113,937                                    | 9.2                     | 32.1  |
| C.P.         | 13,991,863                                       | 468,165                                      | 3.8                     | 33.5  |
| Hyderabad    | 12,855,934                                       | 425,302                                      | 3.5                     | 33.0  |
| Mysore       | 5,970,446  | 314,673                                      | 2.6                     | 52.7  |
| Rajputan...  | 10,330,957                                       | 282,312                                      | 2.3                     | 27.8  |
| Madras       | 42,168,483                                       | 227,184                                      | 1.9                     | 5.4   |
| C.I. Agency  | 7,653,893  | 149,941                                      | 1.2                     | 19.6  |
| Burma        | 5,970,446  | 149,427                                      | 1.2                     | 11.8  |
| Other areas  | 38,477,465                                       | 109,597                                      | 0.9                     | 2.8   |
| Bengal       | 46,109,157                                       | 68,809                                       | 0.6                     | 1.5   |
| Total        | 3,06,406,463                                     | 12,170,439                                   |                         |   |

### GENERAL CHARACTERISTICS OF THE EPIDEMIC WAVES

A typical epidemic wave is generally symmetrical differing with normal curve in regard to the height of the peak. In the pre-epidemic periods it takes some time to develop epizootic among the field and sewer rats (*R. norvegicus*) which may be of sylvatic origin. It takes about 10 days for the epizootic to appear in the house rats and about another 10 days to involve the human beings (figure 3).

The peculiarity of plague epidemics, particularly in cities and towns is that after its introduction or reappearance it goes on for a few years before a time interval is reached, and usually bigger the city the longer is the range of continuity. In a big city like Calcutta where there is enormous rat population it takes several years for a large size epizootic to develop, which is maintained for several years due to almost continuous supply of susceptibles in sufficient number (thus raising the herd immunity) and to bring about cessation. It is also noticed that in the beginning small

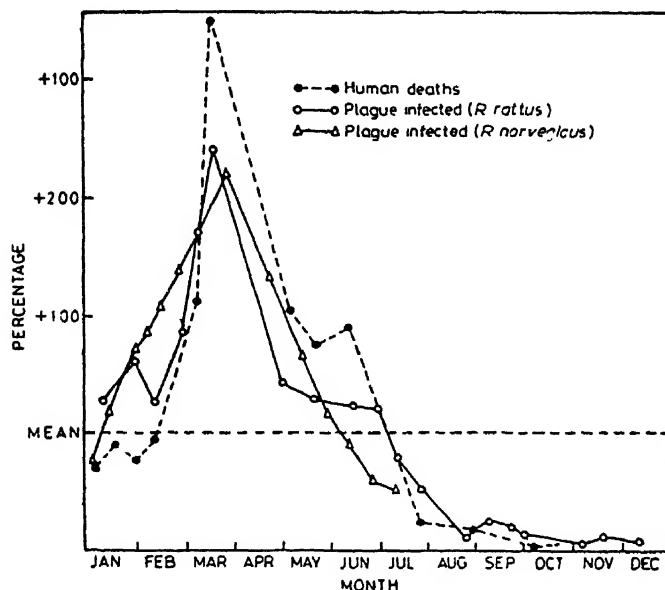


FIGURE 3

outbreaks occur which gain momentum within 3 or 4 years unless seriously interfered within its earlier phase of the first 2 years (as in case of the last Calcutta epidemic—figure 4a) and then a heavy epidemic occurs for some years followed by gradual fall through several years till it disappears for some time (figure 4b). The abrupt decline of epidemic wave as a premature cessation of transmission is considered as "incomplete" epidemic because it terminates before its herd immunity has reached the climax. This is what happened in the last Calcutta epidemic due to prompt and effective early action by the author. The world prevalence of plague at the time of Calcutta epidemic is shown in figure 5.

*Morbidity, mortality and spread:* Cases are generally severe in the early phase of the epidemic. With the progress of the epidemic cases become less and finally are of mild type. All age groups may get the disease except the infants (0-6 months old). Sex and occupation have little significance but in Calcutta men suffered more than the females. M : F ratio may be partly responsible but women are better covered than males. The case fatality in bubonic plague varies in different epidemics, the fatality rate varying between 60 and 90% in untreated cases. It varies with climate, facilities for treatment, overcrowding etc. In the overcrowded Hong Kong and among the Chinese it was 93%, among the Japanese 60%, among the Indians 77% but among the Europeans 18.3%.

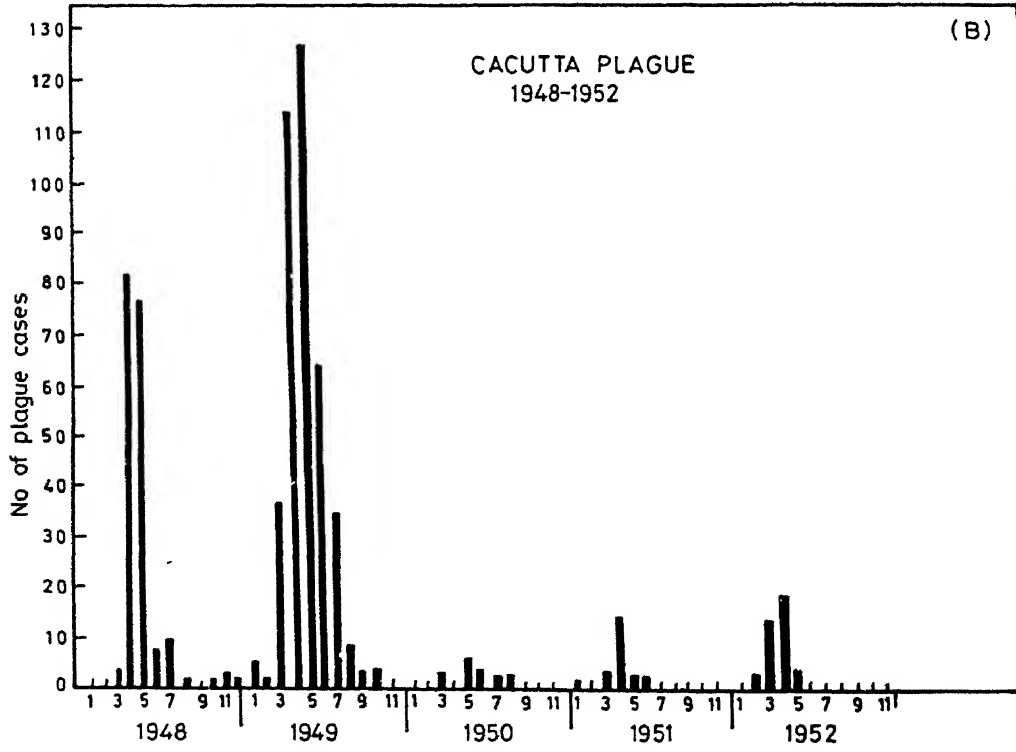
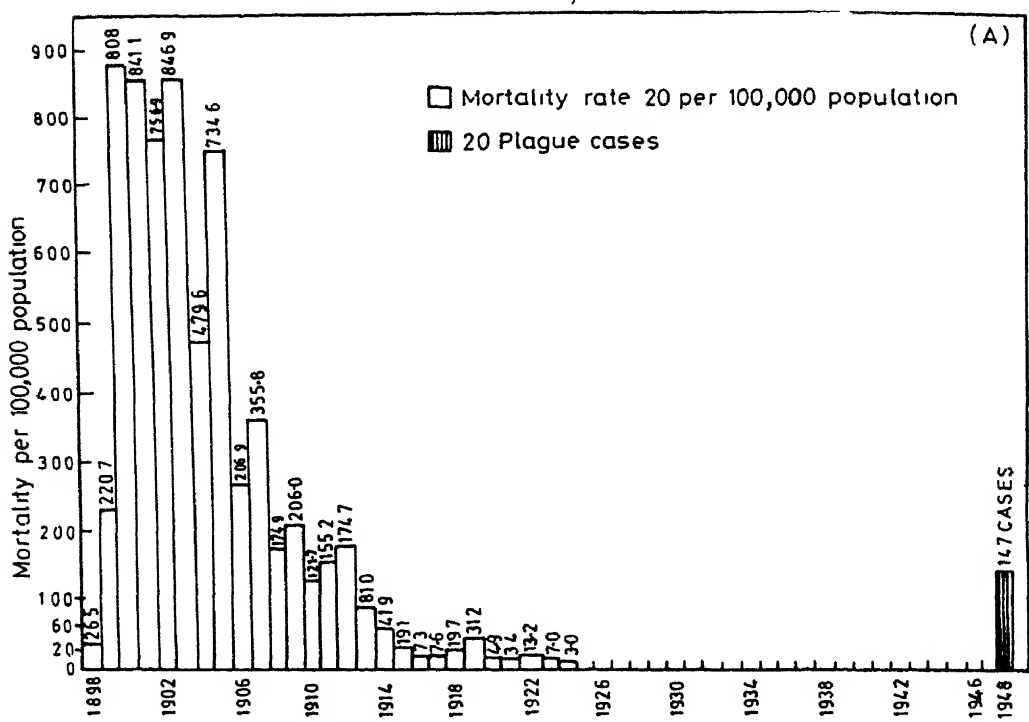


FIGURE 4 (A & B)

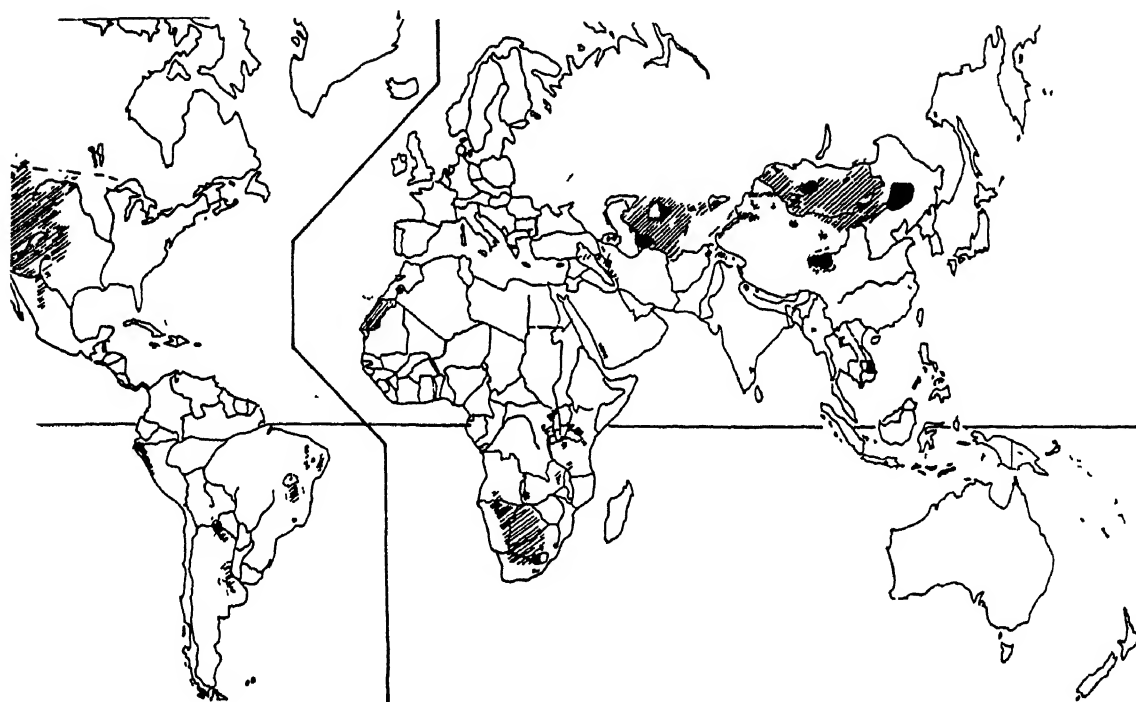


FIG 5 Geographical Distribution of Human Plague reported between 1948 & 1952 and 1957 Areas where Human Cases were Reported

It usually spreads by contiguity but it may also spread to distant places through infected rats or fleas being carried with commodities or in person's body or clothes and also from one country to another through maritime and aerial transport. In a city the first appearance of plague in a locality may be followed by a rise in small foci (as in grain stores, ware houses, etc.) before covering wider areas. It has predilection to dirty insanitary dwellings, more particularly in the ground floor attacking generally poorer class of the people in the rat-infested huts and bustees, granaries, store rooms, refuse dumps harbouring rats.

#### PHASE I

There were only three important discoveries during this phase namely, (i) confirmation of the role of rodents in the transmission of plague, (ii) discovery of plague bacillus in 1894, and (iii) the suggestion of transfer of infection through the maritime routes.

The relationship of rodents to plague actually dates back to the period earlier than 600 BC (vide *Bhagavata purana*). The Bible also mentions about 'mice' whose corresponding Hebrew word was "akbar" meaning rat. In all probabilities mice in the Philistine towns and ships

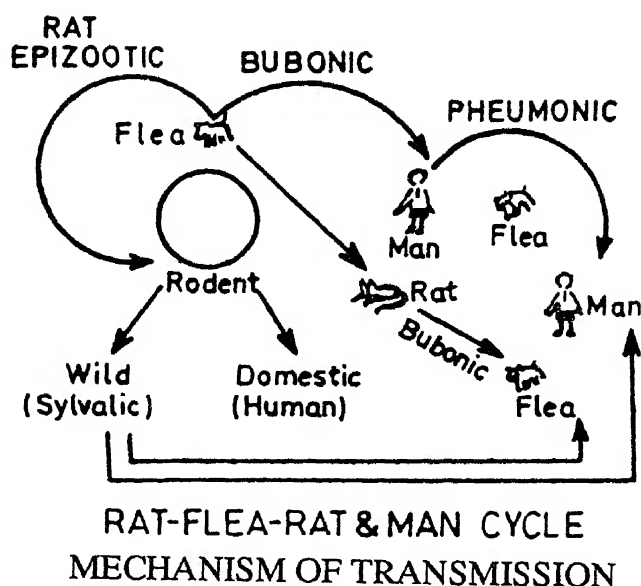
were really rats (*Rattus alexandrianus*). It took a concrete shape in India by the 17th century when in 1616 the treasurer of Jahangir's court described how handling of dead rats led to plague infection in Punjab. A similar relationship was also established in China by Ling Chi (1736-1809). However, upto 1896 it was not known how the infection was actually being transferred from rats to man. It had to remain unsolved till the causative organism was discovered.

Another development during this period was the clinical distinction between bubonic, septicaemic and pneumonic plague. A definite idea about the pneumonic plague is, however, available in the book on '*Manchurian Outbreak*' by Wu-Lien-Teh (1926).

## PHASE II

The actual scientific investigation about plague started when it visited Bombay and began to spread in all directions in India. Three Commissions were engaged in the study of the problems e.g., The German Plague Commission (1899), The Austrian Plague Commission (1895-1900) and the British Indian Plague Commission (1907-1912). The knowledge about the role of rats became clearer by the study of Bombay Plague (1896-1898) by the work of the pioneer French Epidemiologist, Dr Simond who between 1894 and 1898 working independently drew attention to the way plague was spreading in Bombay. By careful experimental device he proved that the infection was transmitted from rats to rat and human being through the bite of infected fleas, the ecto-parasite of rats. As is generally the fate of all new discoveries or theories, this new conception of bubonic plague as a flea-borne infection was opposed by many epidemiologists as well as the Plague Commissions (British Indian and German). However, very soon support of Simond's observation began to pour in from the Japanese, French and Austrian workers. Gauthier and Raybaund (1903) successfully repeated Simond's experiment confirming his findings. Also Glen Liston (1906) in Bombay found evidence of the role of fleas by carrying out experiments with guineapigs which when placed inside the plague-infected house, not only imbibed infection and died but the fleas from their bodies also yielded plague organism. The rat-flea-man cycle was thus discovered as the missing link. This cycle is diagrammatically represented in figure 6a.





Researches indicated that apart from mechanical transmission, which is possible during heavy epidemics, the usual mode of transmission was through blocked fleas as discovered by Bacot and Martin (1914). They showed that infection was generally transmitted by fleas whose gizzard and proventriculus had been blocked by growth of plague bacillus. Under the circumstances the fleas being unable to suck blood into their stomach regurgitated the organisms by elastic recoil of their distended oesophagus and pharynx back into the wound made by their probacis together with masses of plague bacillus detached from the obstructing growth. The blockage phenomenon can be easily seen under microscope or even hand-lens (figure 6b).

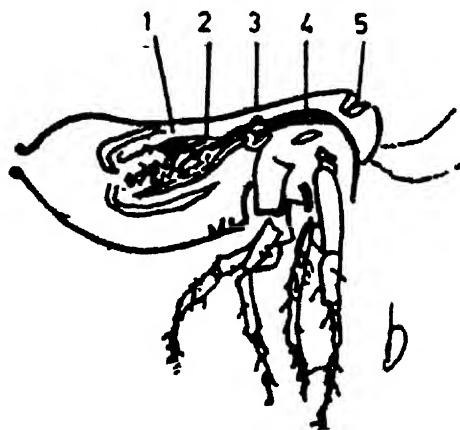


Fig 6B Diagram of blocked rat flea (Bacot & Martin) 1. Masses of *P. pestis* & stomach; 3. Proventricular value blocked; 4. Oesophagus with masses of *P. pestis*; 5. Pharyngeal pump

The next stage was to identify various rodents involved in the transmission of plague infection namely: (i) wild rodents, (ii) peri-domestic rodents, and (iii) domestic rodents.

Some of the wild rodents are tarabagan, susliks, jerboa, gebrils, wood rats, ground squirrels, striped mice, tatera etc. and the peri-domestic rats are *Tatera indica*, *Funumbulus palmarum*, *Mus buduga*, *Gunomys kok* (field rats), while domestic rats are *Rattus rattus*, *Rattus norvegicus*, *Rattus rattus alexandrianus*, *Rattus frugivorus*, *B. bengalensis* etc. (figure 7).

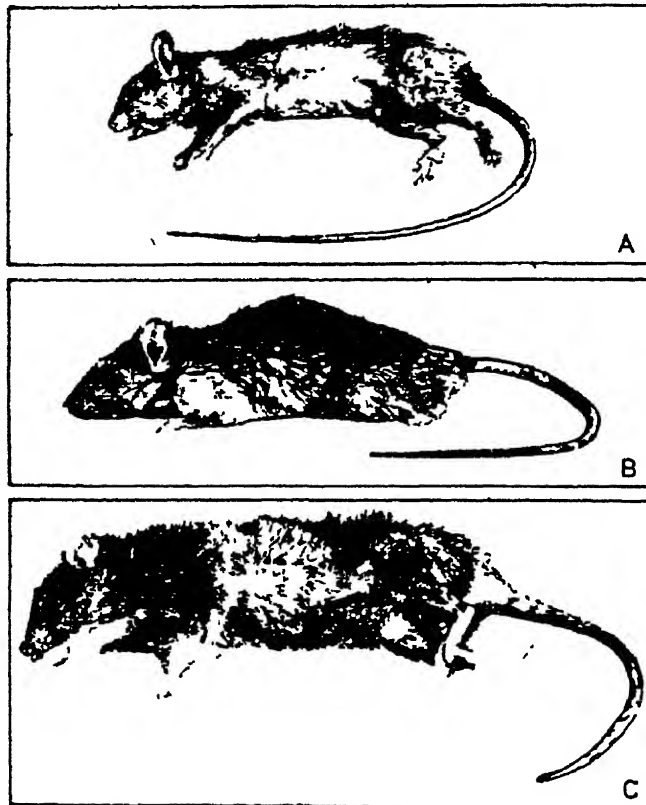


Fig 7 A-C A. *Rattus rattus rattus*; B. *Rattus rattus norvegicus* C. *Bandicota bengalensis* Kok (*Gunomys Kok Auct*)

The above is the nutshell of the position about the knowledge of plague till 1940 when the author took up the study. On taking over he found the position as follows:

There was hardly any method of treatment or clear-cut quick method of identification of plague bacillus, culture from both liquid and solid media being rough and auto-agglutinable. Also knowledge about specific chemical antigenic structure remained vague nor the antigen could be isolated in pure form. The bacillus was taken as normally rough

ignoring the fact that roughness is only a dissociant form of bacteria in general and antigenically poorer than that of the smooth organism. So the use of this organism yielded but poor vaccine and also poor antiserum which was the only effective method of treatment before the discovery of sulphur drugs and streptomycin. The author therefore concluded that the fundamental defect lay in the culture media and was virtually a problem of nutrition of plague bacillus. He also noted that due to the lack of this knowledge, for growing the organism in the media used required a heavy inoculum and 4-6 hr lag period. What was happening is that the media containing half-digested protein of the meat was not immediately suitable for the growth and nutrition of the plague bacillus. The heavy inoculum resulted in the death of a portion of the bacterial mass and this provided the proper nutrition for their growth and elaboration of the enzyme for the digestion of the protein to convert it into amino-acid form needed for their growth and for which the lag period was necessary.

### DISCOVERY OF A NEW MEDIUM

Having reached at this conclusion the primary need was to work out an appropriate and suitable nutritive medium which would provide the ready made nutritive substance for the plague bacillus to grow without any lag period. Thus his discovery of the protein-free casein hydrolysate medium in connection with the culture of diphtheria organism came very handy for the purpose by enrichment with a very small quantity of liver extract. In this medium growth starts with the smallest inoculum (even as low as 10 or 20 organisms) and a count made every half hours shows that the optimum growth of 193 million/ml was reached in 54 hr (table 2) and that this optimum growth could be obtained with hydrolysate of 72 hr digestion. The final preparation of the medium was as follows:— (Seal 1943, Seal & Mukherjee 1950).

*Liquid Medium:* Casein is hydrolysed by combined HCl and  $\text{H}_2\text{SO}_4$  by digesting it for 72 hr. It is then neutralised by  $\text{Ba}(\text{OH})_2$ ,  $8\text{H}_2\text{O}$  and filtered, and nitrogen content estimated by micro-Kjeldahl method. To the filtrate 0.2 mg of anhydrous  $\text{Na}_2\text{HPO}_4$  and 0.067 mg of  $\text{KH}_2\text{PO}_4$  are added per mg of N, followed by 0.04 mg of  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ . The pH is adjusted between 7.4 and 7.6 with NaOH and the mixture heated for 5 min. The precipitate thus formed is filtered off through paper and the filtrate diluted to contain 2.7 mg N/ml. Then  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  and  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  and liver extract are added in doses of 0.5, 0.0004 and 0.1 mg, respectively, per ml of the medium and the pH readjusted, if necessary. The medium is then

distributed in flasks and tubes and sterilised in the autoclave at 110°C for 10 min.

**Table 2**  
*Viable counts/ml of growth of Y. pestis in casein hydrolysate broth medium in relation to the period of incubation*

| Period of incubation (hr) | C H basal | C H with liver extract | Period of incubation (hr) | C H basal            | C H with liver extract |
|---------------------------|-----------|------------------------|---------------------------|----------------------|------------------------|
| (Seed)                    | 556       | 420                    | 18                        | 28x10 <sup>2</sup>   | 266x10 <sup>3</sup>    |
| 1/2                       | —         | 595                    | 24                        | 155x10 <sup>3</sup>  | 286x10 <sup>4</sup>    |
| 1                         | 858       | 645                    | 30                        | 85x10 <sup>4</sup>   | 25x10 <sup>6</sup>     |
| 2                         | 612       | 1,005                  | 36                        | 77x10 <sup>5</sup>   | 65x10 <sup>6</sup>     |
| 3                         | 750       | not done               | 42                        | 872x10 <sup>5</sup>  | 106x10 <sup>6</sup>    |
| 6                         | 2,800     | 4,450                  | 48                        | 98x10 <sup>5</sup>   | not done               |
| 9                         | 4,400     | not done               | 54                        | 1075x10 <sup>5</sup> | 193x10 <sup>6</sup>    |

#### CH, Casein hydrolysate

The advantages of this medium are:

- (i) Autoagglutination is completely removed;
- (ii) Facilitates isolation of specific soluble proteins in pure form by precipitation with sodium or ammonium sulphate;
- (iii) Facilitates production of improved quality of Haffkine plague vaccine and of antiplague horse serum for treatment of plague cases;
- (iv) Permits study of the chemical antigenic structure of the organism;
- (v) Solid medium in the form of casein hydrolysate agar gives smooth colonies and produces a smooth suspension for agglutination test and study of specific proteins and polysaccharide of the plague and allied bacilli;
- (vi) Antisera produced against different fractions of specific proteins helps in differentiating *Y. pestis* from *P. pseudotuberculosis* and other organisms;

- (vii) Antigens isolated from supernatant of the growth of the organism in casein hydrolysate medium can be directly used as vaccine, and can be indefinitely maintained as a stock, to be used in any emergencies as well.

*Solid medium for the culture of plague bacillus:* Ordinary nutrient agar gives poor growth of plague bacillus. Sokhey (1939) found 5% rabbit blood agar most suitable for the purpose but the drawback is that vaccine prepared out of growth of plague organism will also contain some antigen of rabbit's blood which may give rise to adverse antigenic reaction apart from the large quantity of rabbit blood that would be necessary for the production of vaccine on a large scale. Casein hydrolysate agar (Seal 1950) was found to be equally as good as the 5% rabbit blood agar barring the probable undesirable reaction. Besides, plague culture on CH agar can be advantageously used for the preparation of potent antiplague vaccine as well as antisera and also for isolation of specific soluble antigen for serological work namely, agglutination, precipitation and complement-fixation work and estimation of specific antibody against plague infection in both man and rodents on a large scale by gel-precipitation test (to be described later).

### CAPSULE, ENVELOPE AND VIRULENCE

There was no unanimity of opinion about the presence of both capsule and envelope being present in the plague bacillus. The demonstration of both actually depended upon the efficiency of the culture medium as well as on the technique of staining. With the improvement of the culture medium the present author (Seal 1959) was able to demonstrate the existence of both in the virulent plague bacillus. By a new technique *Y. pestis* grown on 5% rabbit blood agar or CH agar at 37°C showed the largest amount of envelope substance and smallest amount in strains grown on A-D agar. The avirulent non-protective plague and pseudotuberculosis strains did not show any envelope (figure 8). The organism is suspended in few ml of distilled water and after 5 min a drop or two of formalin is added to this suspension and then stained with methylene blue by adding a few drops into the suspension. Under the high power microscope envelope is seen as a light-coloured membrane around the capsule stained a little deeply.

### TYPES OF PLAGUE BACILLI

Originally plague bacillus was placed under *Pasteurella* but lately it has been classified in the group called *Yersinia* after its discoverer. Bezsonova



FIG 8A&B. A. Showing the capsule and the core of the organism and an unstained hallow around; B. Showing the envelope and the core of organism differentially stained by a special technique

(1930) divided the strain into two types: (i) Glycerol-positive, and (ii) Glycerol-negative. The strains isolated in south-east Russia, Mongolia, Manchuria and Central Africa were found glycerine-positive, whereas strains isolated in India and south east Asian Islands, and Europe were glycerol-negative and nitrous acid positive as opposed to glycerol positive (fermenting glycerol) and nitrous acid negative. Devignat (1951, 1958) classified the strains into three types. viz., *Var orientalis*, *Var mediaevalis* and *Var antiqua*. Thus, the author felt the needs for a thorough study of the chemical antigenic structure which can place all organisms into a stable and standard classification.

## VIRULENCE OF PLAGUE BACILLUS

Virulence and toxicity of plague bacillus are not synonymous but complementary. It not only means invasiveness but also combined with toxicity it brings about fatal termination of the infected man or animal. Virulence is determined by animal inoculation test. Only 5-10 virulent organisms are required to kill 50% of mice (Sokhey 1939). The newly isolated organisms are tested against this standard for virulence. Plague strains of low virulence have been detected by the author both in the course of an epidemic as well as during the interepidemic period among the rats. Virulence of such organisms can however be enhanced by passage in mice. These two observations taken together provided the clue to the mechanism of plague to rise in epizootic form from the pre-epizootic to the epizootic stage and then to human epidemic. Accordingly, the author planned and carried out experimental epidemiology to be reported later.

## TOXICITY OF DIFFERENT TYPES OF PLAGUE VACCINES

One of the objectionable features of the Haffkine plague vaccine was its toxicity. Once it was even thought that the response to immunity depended upon this reaction but the comparative tests of toxicity as well as of the protective quality of different vaccines carried out by the author (table 3) indicate that the vaccine prepared in casein hydrolysate broth was least toxic and yet gives better immunity.

From table 3 it is apparent that the growth in CH Agar medium is the least toxic and better protective than the growth in casein hydrolysate broth and Haffkine plague vaccine which was found to be the most toxic and comparatively least protective. So the advantages lie with the hydrolysate media. A comparative colony characteristics in different solid media are shown in figure 9 (Seal 1950.).

**Table 3**

*Toxicity and mouse protective doses of three different plague vaccines*

| Types of vaccine  | Mouse-protective dose | Toxic dose |
|---|-----------------------|------------|
| Haffkine plague vaccine   | 0.006 ml              | 0.2 ml     |
| Casein hydrolysate broth (1000 million organisms from 3 days' growth) | 0.0053 ml             | 0.8 ml     |
| C H agar vaccine (1000 million organisms per ml)                      | 0.0048 ml             | 1.5 ml     |

The growth on A-D agar was not only poor but also appeared more rough than the serum or rabbit blood agar and formed unstable suspension in 0.85 salt sol. The growth on rabbit blood agar and C H agar appeared as smooth colonies and formed homogeneous and stable suspension in normal salt solution. Virulent *Y.pestis* also showed some dew drop like colonies on these two media. The colonies of the avirulent non-protective plague and pseudotuberculosis strains were comparatively more opaque with occasional smooth colonies. The special advantage of rabbit blood agar is its capacity to maintain virulence and other antigenic structure in stab culture in Frigidaire with only one subculture annually. While the growth of virulent plague bacillus gives a uniform turbidity in casein hydrolysate broth, the avirulent plague strains give somewhat granular turbidity. The classical stalactite growth or the pellicular growth leaving the media more or less clear, and deposit at the bottom which, according to the author is nothing but evidence of roughness and lack of proper nutrition (figure 10).

#### REACTION ON CARBOHYDRATES (SUGARS)

The author also brought about improvement in the carbohydrate fermentation reaction employing 19 sugars for 35 strains of both plague and pseudotuberculosis organisms (Seal 1951 a). The results are shown in table 4.

**Table 4**  
*Differential sugar fermentation reaction by different strains*

| Strains                                     | Sugars                 |          |                            |            |   |         |         |          |
|---|------------------------|----------|----------------------------|------------|---|---------|---------|----------|
|   | Maltose<br>Isodulcitol | Rhamnose | Raffinose                  | Sachharose | Trehalose   | Insulin | Dextrin | Sorbitol |
| Plague Strains                              | A                      | 0        | 0                          | 0          | A   | 0       | A       | 0        |
| PR/I  | A                      | A        | 0                          | 0          | 0   | 0       | A       | 0        |
| PR/II                                       | A                      | A        | 0                          | 0          | A   | 0       | 0       | 0        |
| PR/III                                      | A                      | A        | A                          | 0          | 0   | 0       | —       | 0        |
| PR/IV                                       | A                      | A        | A                          | 0          | A   | 0       | —       | 0        |
| <i>P. avisepticus</i>                       | 0                      | 0        | 0                          | A          | 0   | A       | 0       | A        |
| 0, not fermented, A, fermented; —, not done |                        |          |                            |            |   |         |         |          |
| Proposed subgroup of <i>Y. pestis</i>       |                        |          |                            |            |   |         |         |          |
| Proposed subgroup                           | Test with glycerol     |          | Production of nitrous acid |            | Remarks   |         |         |          |
| <i>Y. pestis</i> var <i>orientalis</i>      | —                      |          | +                          |            | Oceanic race (Berlin & Borsekov 1938)                 |         |         |          |
| <i>Y. pestis</i> var <i>antique</i>         | +                      |          | +                          |            | Found in Central and North Asia and in Central Africa |         |         |          |
| <i>Y. pestis</i> var <i>mediavalis</i>      | +                      |          | +                          |            | Found in South-East Russia                            |         |         |          |



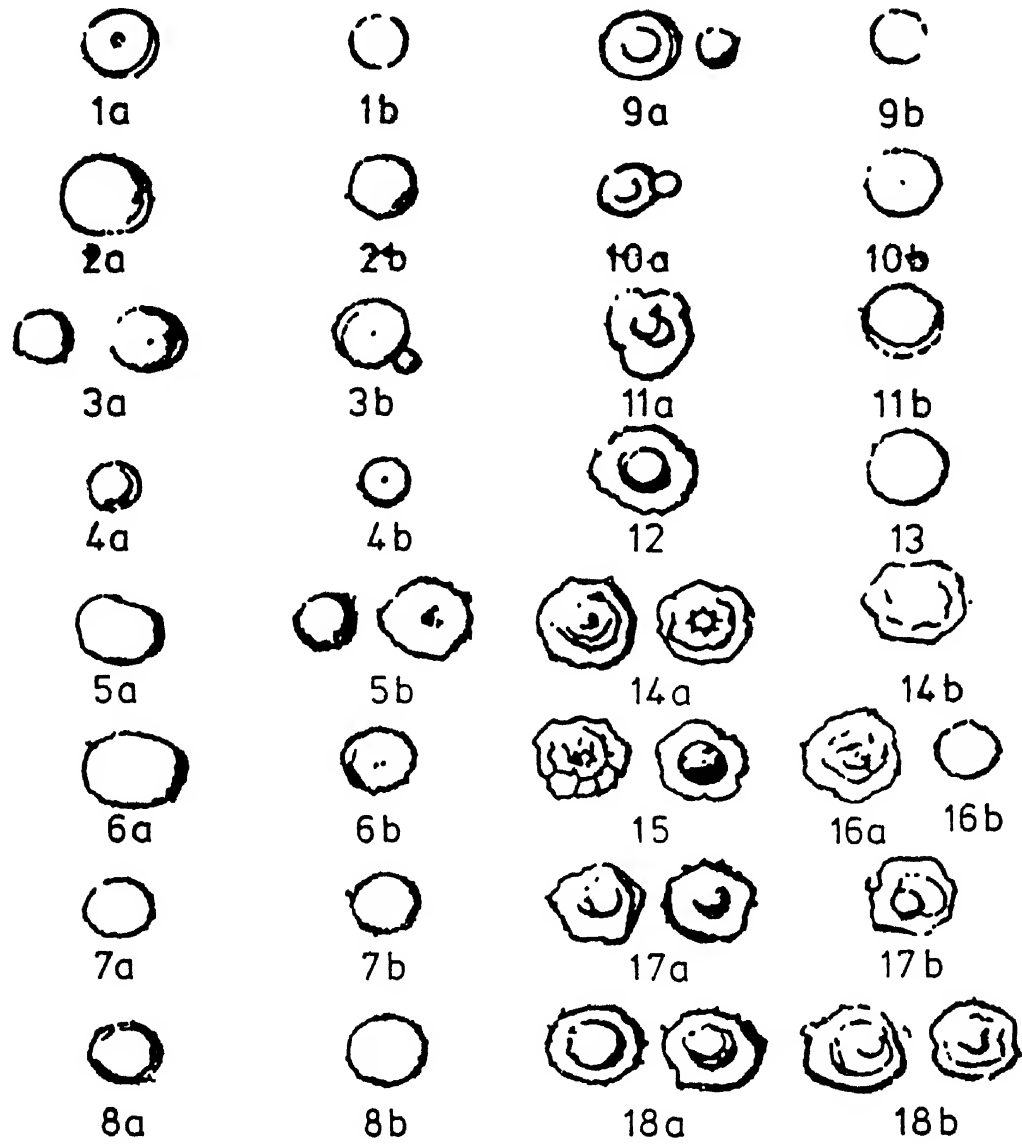


FIG 9 Colonial characters of the various plague and pseudotuberculosis strains on 10% serum and 51 rabbit-blood agar media.

|                           |                   |                  |        |          |       |  |      |
|---------------------------|-------------------|------------------|--------|----------|-------|--|------|
| 1.(a)                     | Virulent          | <i>P. pestis</i> | 14/L   | on serum | agar, | (b) the same on rabbit-blood   | agar |
| 2.(a)                     | "                 | "                | 27/L   | "        | "     | (b) " "  | "    |
| 3.(a)                     | "                 | "                | 149/B  | "        | "     | (b) " "  | "    |
| 4 (a)                     | "                 | "                | 308/B  | "        | "     | (b) " "  | "    |
| 5 (a)                     | "                 | (old)            | 54/H   | "        | "     | (b) " "  | "    |
| 6 (a)                     | "                 | "                | 337/L  | "        | "     | (b) " "  | "    |
| 7 (a)                     | "                 | "                | 139/L  | "        | "     | (b) " "  | "    |
| 8.(a)                     | Av protective     | <i>P. Pestis</i> | 53 Har | "        | "     | (b) " "  | "    |
| 9.(a)                     | "                 | "                | EV     | "        | "     | (b) " "  | "    |
| 10.(a)                    | "                 | "                | Tjs    | "        | "     | (b) " "  | "    |
| 11.(a)                    | Av non protective | "                | 120/5  | "        | "     | (b) " "  | "    |
| 12.(a)                    | "                 | "                | Har    | "        | "     | 13. TJR on serum agar<br>17 & 18 Pseudotubercule<br>Strains PRI and PR<br>(II) |      |
| 14.15, 2nd 16 non-protect |                   | av.p pestis      | 1/av   | "        | "     |  |      |

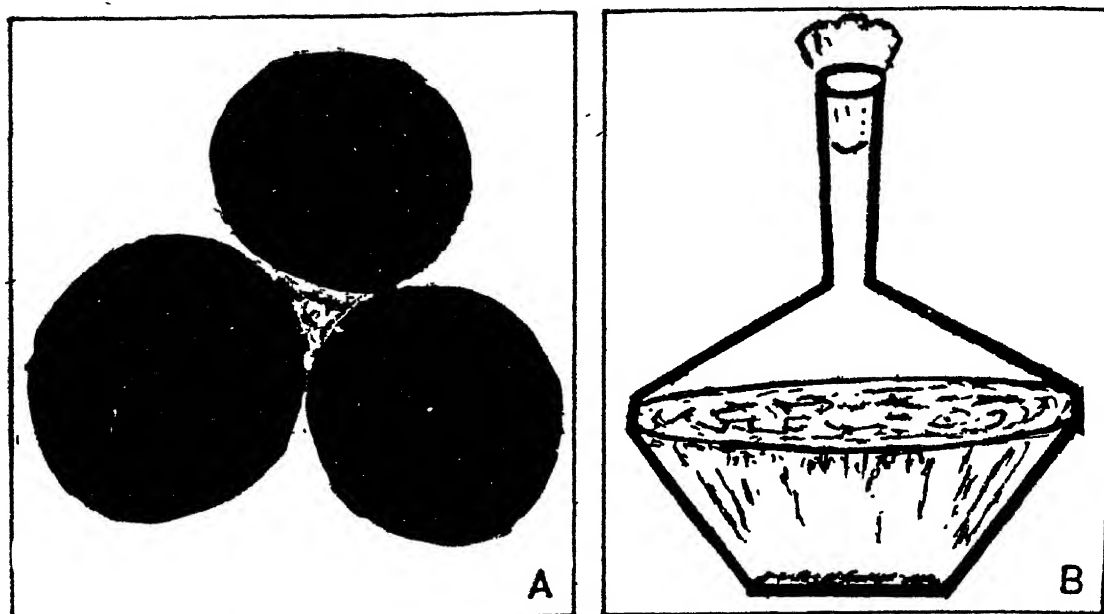


FIG 10 A&B A. Smooth Colonies of Virulent plague strains, B. Culture of plague bacillus in Haffkine Flash

The sugars that distinguish pseudotuberculosis from the plague strains are rhamnose and iso-dulcitol producing acid with the former and not with the latter strains. *P. avisepicus* did not ferment maltose and rhamnose but fermented sachharose and sorbitol, thus differentiating it from both plague and pseudotuberculosis organisms. To distinguish pseudotuberculosis strains from one another the three sugars viz., raffinose, trehalose and dextrin were found useful, dextrin being fermented by PR/I, trehalose by PR/II, raffinose by PR/III and both raffinose and trehalose by PR/IV. (PR - *Pseudotuberculosis rodentium*).

#### IDENTIFICATION OF CHEMICAL ANTIGENIC STRUCTURE OF PLAGUE AND OTHER STRAINS

As mentioned earlier the plague workers were facing difficulties in the identification of the organism due to autoagglutination and in isolating the specific antigen, as the media contained non-specific proteins coming in the way for purification of the antigen. Both the difficulties were overcome by the discovery of the protein-free casein hydrolysate medium. A third innovation was introduced in the agglutination reaction by using live plague culture instead of killed or chemically treated suspension. The problem actually lay in the differentiation between plague and pseudotuberculosis strains, both of them being the carrier and infecting

organism for the rats. Another difficulty faced by them is the differentiation between virulent and avirulent protective and non-protective strains, all the three having been found by the author to play some role in the epidemiology of plague. Burrows and Bacon (1956) by a special technique discovered two additional antigens called V and W which can differentiate virulent from the avirulent plague strains being present in the former and absent in the latter. In a subsequent study Burrows (1959) discovered two additional determinants of virulence namely P (pigment-positive) and Pu (Purine producing). Accordingly the serologic antigenic structure of plague and pseudotuberculosis strains were defined as follows (table 5).

**Table 5**  
*Antigenic structure of plague bacilli as defined by Burrows (1959)*

| Strains                                  | Antigens               | Antigenic character                                    |
|--|------------------------|--|
| <i>Y Pestis</i> virulent                 | F1+, VW+, P+, Pu+      | Common rough somatic antigen                           |
| <i>Y pestis</i> avirulent protective     | F1+, VW-, P+ or -Pu+   | --do--   |
| <i>Y pestis</i> avirulent non-protective | F1-, VW-, P+ or -, Pu- | -do  |
| <i>P pseudotuberculosis</i> (fresh)      | F1-, VW+, P+, Pu+      | Flageller, smooth somatic<br>(type and group specific) |
| <i>P pseudotuberculosis</i> (laboratory) | F1-, VW-, P+, Pu+      | Common rough somatic                                   |

F1 antigen of Baker et al is the same as antigen A of Seal (1943, 1950, 1951 d, 1953, 1954)

## CHEMICAL ANTIGENIC STRUCTURE OF PLAGUE AND ALLIED ORGANISMS

Various workers from time to time attempted to isolate specific antigens or antigens in pure form (Schutze 1932, Sokhey & Maurice 1935, Sokhey 1939, and Baker et al. 1947) all faced the difficulty of freeing it from the non-specific part mixed up with it although they noted that it was a protein or a nucleoprotein. Baker's work also suffered from the defect in choosing nutrient agar which was not fully nutritive to plague organism. Author's casein hydrolysate medium seemed to be the ideal in this respect (Seal 1951b).

## TECHNIQUE OF ISOLATION IN BRIEF

From *casein hydrolysate broth*: Casein hydrolysate broth was inoculated with virulent plague strain grown on either 5% rabbit blood agar, casein hydrolysate broth or casein hydrolysate agar for 48 hr at 37°C, and

incubated at 37°C for 2-3 weeks and filtered. The clear filtrate is then precipitated at different saturation of  $\text{Na}_2\text{SO}_4$  to yield P1/3, P1/2; P1/2-1/3 and P<sub>1</sub>-1/2 fractions.

- (b) *From casein hydrolysate agar:* 72-hr culture of *Y. pestis* virulent was inoculated into the CH agar in Roux bottles and incubated at 37°C for 48 to 72 hr and washed with distilled water and collected in a flask in strength of an optimum of 600 million organisms per ml (by matching with opacity tubes) and incubated for 12-14 days by which period all suspended organisms would be found dissolved (culture negative). The material is then filtered and the filtrate precipitated by  $\text{Na}_2\text{SO}_4$  in the same saturation of  $\text{Na}_2\text{SO}_4$  to yield the same four fractions as in the case of casein hydrolysate broth.

Chemically tested, all the protein fractions were sulphur phospho-proteins. Tests for tryptophane and tryosine showed that the P1/2 fraction obtained from the protective strains grown in CH broth contains both these amino-acids and those from non-protective plague and pseudotuberculosis strains showed little or none, whereas those from the water-soluble extracts did not contain any of them. Since tryptophane was virtually absent in the CH broth its presence in the fractions apparently indicate that the plague bacillus could synthesize this amino-acid during its growth in CH broth.

### ISOLATION OF ACTIVE POLYSACCHARIDE FROM PLAGUE BACILLUS

In the study of chemical antigenic structure it was of interest to know if any polysaccharide moiety was involved in the specific soluble protein antigen, as we had in case of *V. cholerae*. Attempts were therefore made to isolate this substance from three sources e.g., (i) supernatant of antiplague vaccine; prepared on casein hydrolysate broth; (ii) the bacterial debris of the same vaccine; (iii) specific soluble protein of plague bacillus.

A polysachharide corresponding to "arabinose" yielding osazone with melting point of 166-168°C was isolated more or less from each of the above. It was absent in the non-protective plague and pseudotuberculosis strains. Therefore, the specific soluble protein of plague bacillus is a polysaccharide-protein complex, i.e. a nucleo-protein derivative (Seal 1951 c).

The results of agglutination and cross-agglutination titres of the anti-protein sera against various plague and allied organisms (table 6)

show  $L_{Auto}P1/3^*$  is highly specific plague bacillus so far as the serological reactions are concerned. This fraction is not present in the precipitate of pseudotuberculosis organism or in any of its fractions. Thus  $L_{Auto}P1/3^*$  fraction may be suitably utilised for differentiating protective from non-protective or pseudotuberculosis strains but not the virulent from the avirulent protective plague strains. On the other hand, the fraction  $P1/2-1/3^*$  of virulent plague strain reacting with the plague and pseudotuberculosis strains may be the antigen that maintain the antigenic link between the virulent plague and pseudotuberculosis organisms. It is also related to 'O' antigen of the virulent plague bacillus as seen in table 7.

Table 6

*Agglutination and crossagglutination titres of the anti-protein sera against various plague and allied organisms (min. dil. 1:25)*

| Anti-protein     | Organism (suspensions) |        |         |       |      |       |
|------------------|------------------------|--------|---------|-------|------|-------|
| sera             | 337/L                  | 53H/av | 120H/av | TRU   | PR/I | PR/IV |
| $L_{auto}P1/3$   | 400                    | 200    | 0       | 0     | 0    | 0     |
| $L_5P1/3$        | 200                    | 200    | 25      | 0     | --   | --    |
| $L_5P1/2-1/3$    | 0                      | 0      | 50      | 25    | 25   | 25    |
| $L_5P1/2$        | 100                    | 100    | 25      | 25    | 25   | 25    |
| $53H/avP1/3$     | 100                    | 100    | 25(f)   | 0     | 0    | 0     |
| $53H/avP1/2-1/3$ | 0                      | 0      | 100     | 25    | 25   | 25    |
| $120H/avP1/3$    | 0                      | 0      | 100     | 100   | 25   | 25    |
| $120H/av1/2-1/3$ | 0                      | 0      | 100     | $\pm$ | 25   | 50    |
| TRUP1/3          | 0                      | 0      | 0       | 50    | 25   | 25    |
| PR/IP1/3         | 0                      | 0      | 25      | 50    | 50   | 25    |
| PR/IP1/2-1/3     | 0                      | 0      | 25      | 50    | 50   | 50    |
| PR/IVP1/3        | 0                      | 0      | 25      | $\pm$ | 25   | 50    |
| PR/IVP1/2-1/3    | 0                      | 0      | 50      | $\pm$ | 50   | 50    |

337/L, vir. plague; 53H/av, av. protect. plague; 120H/av. av non-prot. plague PR/I and PR/IV, pseudotub. organism P1/3, 33% saturation of  $Na_2SO_4$ : P1/2, 50% saturation of  $Na_2SO_4$

\*  $A_{Auto}$ : L, Lot; Auto, Autolysate

Table 7

*Result of agglutination test between the virulent plague strain 337/L boiled and the pseudotuberculosis strains*

| Antiserum against<br>337/L 'O' (boiled) | Pseudotuberculosis strains |      |      |         |
|---|----------------------------|------|------|---------|
|   | 1104/B                     | 3572 | 3570 | 1093/25 |
| Result                                  | ++                         | ++   | ++   | ++      |

Thus the relationship of plague to the pseudotuberculosis strains is through the somatic antigen represented by the P1/2-1/3 fraction of broth filtrate of virulent plague bacillus.

Accordingly, the fraction P1/3 of virulent plague bacillus has been designated as Antigen A and P1/2-1/3 as Antigen B, the latter serving as the link between the plague and pseudotuberculosis organisms.

The other advantages of the specific soluble proteins are: precipitation test, ring precipitation test and complement-fixation test, all of which can specifically identify the plague organism. While the gel-precipitation test is to test immunity status of rat population on a mass scale in the field, and lastly for standardisation of anti-plague sera by quantitative flocculation test. A brief reference to these tests may only be made.

### PROCEDURES FOR PRECIPITATION REACTIONS

For direct and cross-precipitation test the specific protein, antigen A, diluted to contain 0.1 mg/ml (1:10000) is distributed in Felix tubes in 0.5 ml amounts. The tubes are lightly shaken and incubated at 37°C for 2 hr. A preliminary reading is taken and the tubes are placed in the refrigerator for the final reading next morning. The highest dilution showing the visible precipitation is taken to be the titre of the serum.

### RING PRECIPITATION TEST

For quicker diagnosis, ring precipitation test may also be done in the following manner: Antigen A 1/1000 dilution is taken in Dryer's tube and the patient's undiluted serum or 1:2 to 1:10 dilution is added on the top of the antigen by means of a capillary pipette. A precipitating ring develops in 1/2 hr. at room temperature. Even the filtrate of casein hydrolysate broth inoculated with unknown organism or the water extractable filtrate

of distilled water suspension of the unknown organism may be used for both ring and tube precipitation tests.

The precipitation titres of various antisera as obtained against L<sub>2</sub>P<sub>1/2</sub> fraction (antigen A) are given in table 8.

**Table 8**

*Precipitation titres of various antisera against L<sub>2</sub>P<sub>1/2</sub> protein fraction*

| Antisera        | Live antigen used for immunisation  | Titre of L <sub>2</sub> P <sub>1/2</sub> fraction |
|-----------------|---|---|
| 54/H rabbit     | virulent <i>Y. pestis</i>   | 2,000,000   |
| H/23 of 2.11.41 | virulent <i>Y. pestis</i> , vaccine filtrate and relatively avirulent E V. strain | 1,000,000   |
| H/27 of 2.11.41 | virulent <i>Y. pestis</i> , vaccine filtrate and relatively avirulent strain Tjs  | 1,000,000   |
| H/45 of 2.11.41 | Relatively avirulent Tjs only   | 1,000,000   |
| H/44 of 2.11.41 | Relatively avirulent Tjs only   | 500,000   |
| E V. Rabbit     | Relatively avirulent <i>Y. pestis</i> E.V. Strain                                 | 500,000   |
| Tjs Rabbit      | Avirulent protective <i>Y. pestis</i>   | 500,000   |
| Tjs Rabbit      | Avirulent <i>Y. pestis</i> Rough  | 200,000   |
| TRU             | Avirulent non-protective <i>Y. pestis</i>   | Cloudy only                                       |
| PR/I Rabbit     | Pseudotuberculosis strain   | --do--  |
| PR/II Rabbit    | --do--  | --do--  |

The result described above thus correspond exactly with those of agglutination test.

### ANTIGENIC CLASSIFICATION OF PLAGUE STRAINS

On the basis of the above analysis the plague strains were classified into three groups namely

- (i) those containing both A and B antigens e.g. all virulent and relatively avirulent strains;

- (ii) those containing a small amount of residual antigen A and full antigen B e.g. relatively avirulent TJS strain;
- (iii) Those containing only antigen B, e.g. avirulent non-protective plague organism like TRU and all pseudotuberculosis strains.

### SPECTROGRAPHIC ANALYSIS OF THE PROTEIN FRACTIONS

A difference between virulent protective and avirulent protective plague strains could be made only by spectrographic analysis of the fractions. The absorption curve obtained by plotting the extinction coefficient against wave lengths the absorption range of all the protein solutions (made by dissolving in N/80 NaOH to make 0.04 per cent clear solution) lay near the ultra-violet region between 230 and 430 m (figure 11). In the case of protein fraction  $L_3$  P1/3 and  $L_{Auto}$ P1/3 fractions isolated from virulent *Y. pestis* had a moderate bend in the absorption curve between 268 and 305m which was not seen with the other preparations. This might be taken as evidence of difference between virulent and avirulent plague strains (Seal 1960a).

### COMPLEMENT-FIXATION TEST

The author (1953) utilised specific soluble plague proteins for complement fixation test. No cross-reaction was noted with antisera against either the whole of avirulent non-protective strains or of the pseudotuberculosis strains.

*Technique of the test:* Antigen A (P1/2 fraction) is diluted to 1:4000 to 1:5000 in 1% NaCl solution. The test sera are inactivated by heating in a water bath at 54°C for 10 minutes. Fresh complement is obtained by bleeding guineapigs each day and is used after titration as in the usual case of Wasserman Test. *Titration:* Tubes are placed in a series and 0.25 ml of various dilutions (1:4 to 1:200) of inactivated sera is added to them, followed by 0.25 ml of 1-1/2 to 2 MHD of complement. The contents are then mixed well by shaking and the tubes incubated at  $0 \pm 3^\circ\text{C}$  for 1/2 hour. Sensitized cells are then added in 0.25 ml amounts and the tubes are shaken and incubated for half an hour at 37°C. A preliminary reading is taken and the tubes are left in the refrigerator overnight and the final reading is taken next morning. The positive results are recorded 1+ to 4+ according to the haemolysed cells settled at the bottom. Negative is indicated by complete haemolysis. The results of complement fixation



tests with protein fractions against the various antiplague and anti-pseudotuberculosis sera are given in table 9.

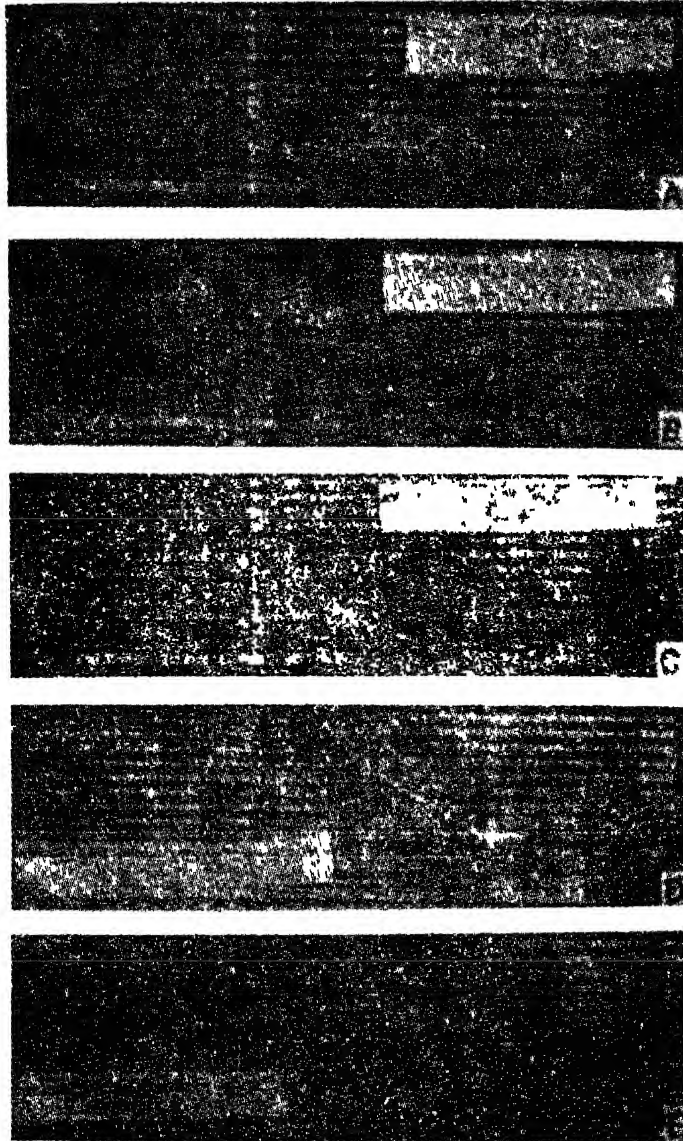


FIG 11 Spectro-photographs of protein fractions of plague and pseudo-tuberculosis strains. A, Virulent protective Plague Strain  $L_{\text{auto}}P1/3$ ; B, Virulent protective Plague strain  $L_5P1/3$ ; C, Virulent protective Plague strain 53 Hav  $P1/3$ ; D, avirulent non-protective Plague strain TRU  $P1/3$ ; and E Pseudotuberculosis strain PR/1  $P1/3$

Table 9

*Complement-fixation tests of 139/L (vir. plague), 53H/av (av, protective plague) TRU (av. non-protective plague), and PR/I (anti-pseudotub.) sera with various protein fractions at dilution 1:5000*

| Antigens                | No             | 139/L |      | 53H/av |      | TRU  |      | PR/I |      |
|-------------------------|----------------|-------|------|--------|------|------|------|------|------|
|                         |                | 1 25  | 1.50 | 1:25   | 1.50 | 1:25 | 1.50 | 1 25 | 1 50 |
| L <sub>5</sub> P1/3     | A <sub>1</sub> | 4+    | 4+   | ++     | 1+   | ±    | 0    | 2+   | 1+   |
| L <sub>5</sub> autoP1/2 | A <sub>2</sub> | 4+    | 3+   | ++     | 1+   | ±    | 0    | 0    | 0    |
| 53H/avP1/3              | A <sub>3</sub> | 4+    | 4+   | 4+     | 3+   | ±    | 0    | 0    | 0    |
| 120H/avP1/3             | B <sub>1</sub> | 2+    | 1+   | 1+     | 0    | 1+   | 0    | 0    | 0    |
| TRU P1/2                | B <sub>2</sub> | ±     | 0    | ±      | 0    | 2+   | 1+   | 1+   | ±    |
| PR/I P1/2               | B <sub>3</sub> | 1+    | 0    | 1+     | 0    | ±    | 0    | 4+   | 3+   |
| PR/IV P1/2              | B <sub>4</sub> | 0     | 0    | 0      | 0    | 0    | 0    | 0    | 0    |

In this test also L<sub>Auto</sub>P1/3 proved to be the most specific followed by other protein fractions which were somewhat broader such as L<sub>5</sub>P1/3 which reacted with the PR/I pseudotuberculosis antiserum. Some workers however opine that this test may be advantageously utilised for the diagnosis of rodent plague in the field especially when the isolation at autopsy is rendered difficult by contamination or decomposition. This is no longer needed since the discovery of gel-precipitation test by the author (1972a), a much simpler method than the complicated complement-fixation test for the diagnosis of wild rodent plague.

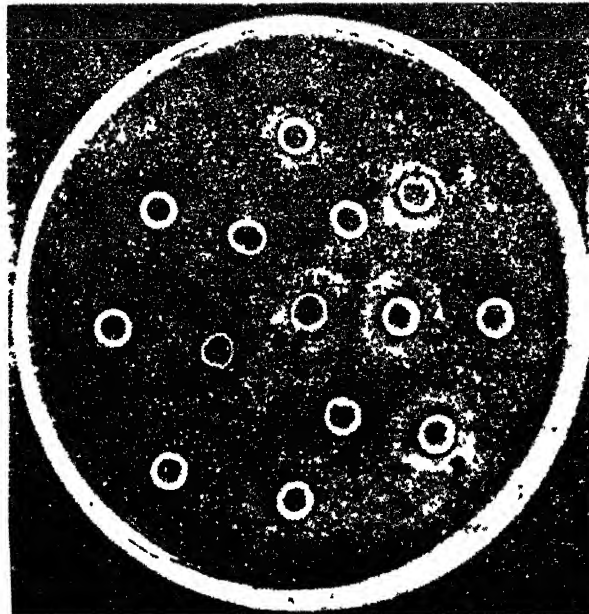
### GEL-PRECIPIATION TEST

*Preparation of the gel plate:* A gel prepared with 1 per cent bacto-agar and 0.1 mg/ml of the antigen A (e.g., L<sub>Auto</sub>P1/3, steamed at boiling temperature of water is mixed with 100 units of penicillin-cum-streptomycin per ml after cooling and poured into petri dish as a thin layer and allowed to solidify. About 16 to 20 holes are cut into the gel. *The test:* Sera collected from the rats and preserved in the refrigerator are separately placed into holes of the plate by using capillary pipettes for each of the dilutions of 1:20 and 1:40. In the central hole a known antiplague serum is placed as positive control and one or two negative controls such as cholera or typhoid antiserum. The Petri dish is then incubated at 37°C for 24 to 48 hours (figure 12). The positive test is indicated by a clear opaque ring

round the well. It is positively useful in surveillance operation in plague control work. Comparative values of the results of immunity test against plague infection in Calcutta and Murarai (West Bengal) in 1971 are shown in table 10.

**Table 10**  
*Comparative values of immunity test in rats against plague infection in Calcutta and Murarai (West Bengal) in 1971*

| Place    | Methods used             | Percentage of rats found resistant |                  | Total |
|----------|--------------------------|------------------------------------|------------------|-------|
|          |                          | <i>B. bengalensis</i>              | <i>B. rattus</i> |       |
| Calcutta | Biological test          | 19.3                               | 37.5             | 20.9  |
|          | Sero-agglutination test  | 32.7                               | 42.1             | 34.0  |
|          | Complement-fixation test | 28.3                               | 36.4             | 29.6  |
|          | Gel-precipitation test   | 30.0                               | 42.1             | 31.9  |
| Murarai  | Biological test          | 19.3                               | 24.8             | 22.0  |
|          | Complement-fixation test | 33.3                               | 25.0             | 23.0  |
|          | Gel-precipitation test   | 23.5                               | 27.5             | 24.4  |



**FIG 12** Gel-precipitation tests: Out of fourteen sera tested, sera Nos 5, 6, 8, 13 and 14 are positive

The above data shows that the immunity test with gel-precipitation test is simpler, quicker and also more reliable.

## SEROLOGICAL METHODS FOR DIAGNOSIS OF PLAGUE

The various methods for sero-diagnosis of plague are agglutination, haemagglutination, precipitation, complement-fixation, and flocculation. The methods of precipitation, complement-fixation and flocculation tests have already been discussed.

## AGGLUTINATION

Although various workers tried to solve this problem of agglutination, it remained unsolved till the author became actively engaged in the work. With the improvement of the growth medium auto-agglutination was completely removed. He made detailed study with antisera raised against virulent *Y. pestis*, *Y. pestis* boiled, for 1/2hr, *Y. pseudotuberculosis*, water-soluble protein extract of *Y. pestis* and with mutually absorbed sera (table 11). He showed that not only the difference between plague and pseudotuberculosis could be distinguished, the inter-strain differences among the plague stains could also be made (Seal 1952).

*Antigen suspension:* In this respect following the author's previous experience with vibrio agglutination (Linton & Seal 1935) by using live culture for preparing the suspension the same principle was applied in case of plague and its allied organism. By using the 24-28 hr growth on 5% rabbit blood agar or CH agar a smooth suspension of the plague organism could be easily obtained and so with pseudotuberculosis and other organisms. Only in case of non-protective avirulent plague strains an additional precaution of filtering the suspension through a pellet of sterile cotton wool was undertaken. For comparative study the various other types of suspensions used were:

- (i) heat-killed organism with or without preservative,
- (ii) boiled organism,
- (iii) alcohol-treated organism and (iv) specific soluble proteins (Seal 1951a).

*Technique of agglutination:* To various dilutions of antiserum placed in uniform bore Felix tubes equal amount of suspension is added and each tube gently shaken. The mixture is then incubated at 37°C overnight and the reading is taken next morning (a preliminary reading

after 2 hours of incubation may also be recorded). The agglutination reaction was of two types viz., (i) floccular or woolly and (ii) granular (figure 13). The former is related to the surface or protective antigen and gives comparatively low titre agglutination, and the latter is related to the somatic antigen which generally gives high titre agglutination. The results of direct and series of cross-agglutination with absorbed and unabsorbed sera (Seal 1951a) gives the serologic relationship between different plague and pseudotuberculosis organisms as shown in table 11.

**Table 11**  
*Serological relationship between *Y. pestis* and *P. pseudotuberculosis**

| Antisera produced against  | <i>Y. pestis</i>      |                          | <i>P. pseudotuberculosis</i> |
|--|-----------------------|--------------------------|------------------------------|
|  | vir. avir. protective | Avirulent non-protective |                              |
| Virulent <i>Y. pestis</i>  | +                     | +                        | +                            |
| Virulent <i>Y. pestis</i> absorbed with <i>P. pseudotuberculosis</i> | +                     | 0                        | 0                            |
| <i>Y. pestis</i> boiled for 30 min                                   | 0                     | +                        | +                            |
| <i>Y. pseudotuberculosis</i>   | 0                     | +                        | +                            |
| Water-extractable protein of <i>Y. pestis</i> (virulent)             | +                     | 0                        | 0                            |
| <i>Y. pseudotuberculosis</i> absorbed with <i>Y. pestis</i>          | 0                     | 0                        | +                            |

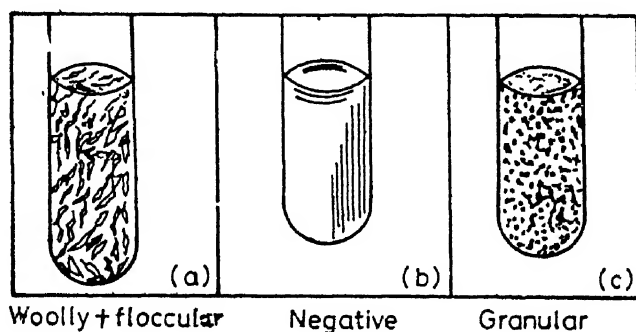


FIGURE 13

In the diagnosis of suspected case of plague or for retrospective diagnosis of human plague a slide agglutination test may help in quick

diagnosis but the tube agglutination described above is the best. Using only No. 1 or 2 of the sera or only No. 5 reliable identification of plague organism can be made.

In the diagnosis of suspected case of plague or for retrospective diagnosis of human plague a slide agglutination test may help in quick diagnosis but the tube agglutination described above is the best. Using only No. 1 or 2 of the sera or only No. 5 reliable identification of plague organism can be made.

Out of 26 bacteriologically positive cases the sera collected between 26 and 130 days after cure 23 showed positive serological reaction and only 3 doubtful reaction while out of 56 suspected cases sera of 5 cases collected between 24 and 134 days after cure showed positive result. As a control, out of 55 non-plague cases 2 were found positive but these two cases gave history of receiving antiplague inoculation. Also by using this agglutination test the various pseudotuberculosis strains were differentiated.

### STEPS TOWARDS ERADICATION OF THE DISEASE

In spite of the knowledge gained about the organism and its specific antigenic structure the main part of the work that still remained unsolved is the knowledge of how the epidemics arise, its periodicity maintained and what happens during the inter-epidemic period. Without this knowledge the question of effective control and final eradication could not be achieved. However, opportunities presented themselves by the reappearance of plague in Calcutta in 1948, the venue of the author's new assignment in the Epidemiology Department of the All India Institute of Hygiene and Public Health, Calcutta.

The last recorded case in Calcutta was in the year 1925. It appeared in 1948 from the same area as in 1895. After the local authorities were convinced that it was an epidemic of plague the author's first task was to find out whether the source of infection was in the city itself or imported from outside. After a preliminary investigation carried out in collaboration with Professor Lal (Lal & Seal 1948) the author rejected the idea of importation of infection from outside. So the important epidemiological knot was the question of origin of infection and how it surfaced after 23 years. There was however one instance of positive isolation of plague bacillus in 1936 from a rat by a research student (Rao 1936) who wanted to study the actual situation underground. The present study also gave us

the clue to the actual mode of spread of infection to human beings. The first human case was detected in Ward 8, the place where the first case of plague appeared in 1895. In the present outbreak dead rats were found in the house of the first case, which the inmates had to handle. Within a few days dead rats were picked up from several other wards of the city and interestingly rats in Ward 8 virtually disappeared. These two situations could be connected to find out the mode of local spread covering half a dozen wards within the span of a week.

### EXPERIMENTAL STUDY ON RAT MOVEMENT

To further confirm the characteristics of rats, experimental studies were conducted on their short and long range movements. Normally the range was found to be not exceeding 50 yards but when they were captured, marked and released 5 to 10 miles away from the place of capture they were found to retrace their paths towards the original place of capture figure 14 Seal & Bhattacharjee 1961, (b).

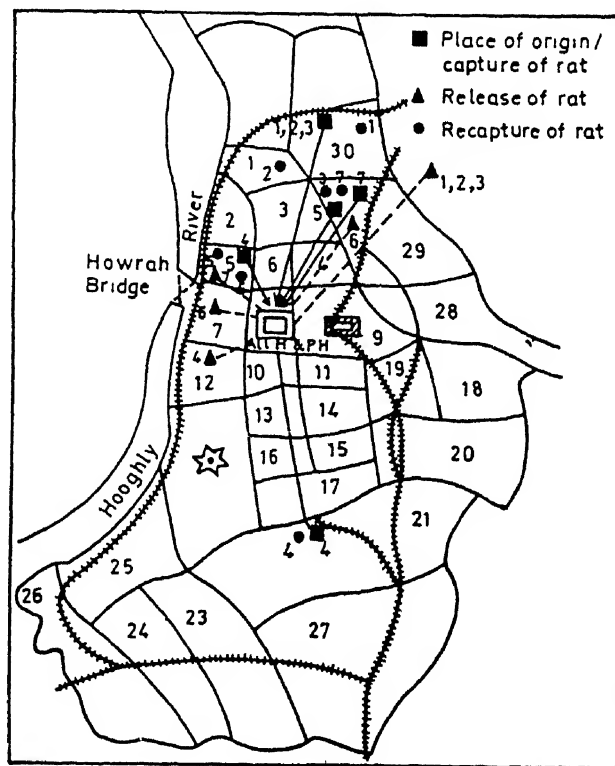


FIG 14 Map of Calcutta

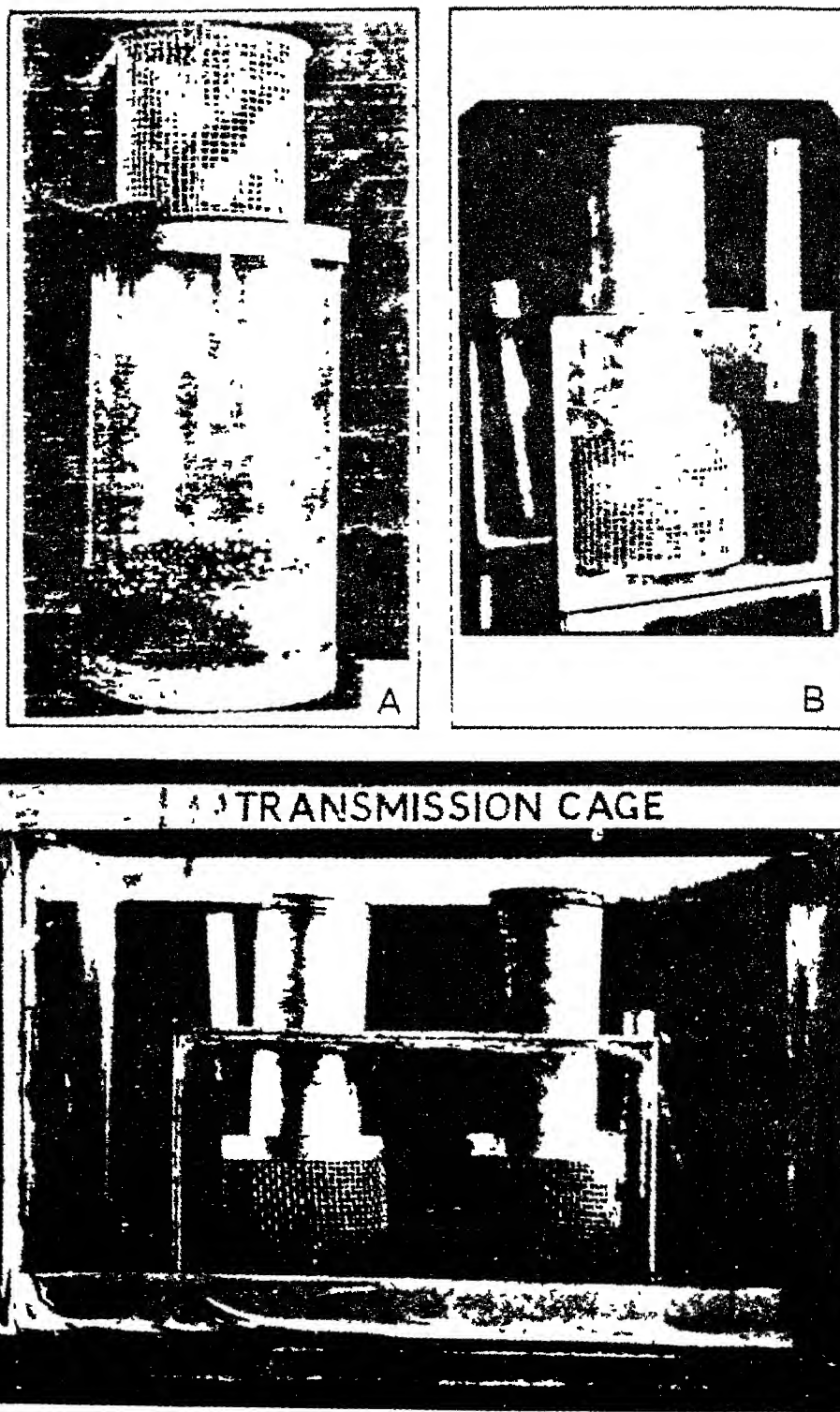


FIG 15 A-C A. Jar for breeding fleas with a rat in a cage for feeding fleas; B. Single cage for transmission experiment; C. Transmission Chamber With double Cage



Two other experiments were organised and carried out to obtain the final clue to the seasonal outbreak, and to the maintenance of carriers of infection for long periods among the rats as well as to study the mechanism by which an epizootic is caused followed by human epidemic.

Results of experimental work on the bionomics of fleas grown in the Laboratory: (Seal & Bhattacharjee 1961a).

The experiments (figure 15) show that while reproduction goes on all the year round in *X. cheopis*, the vector of plague infection has two waves of growth, one in the winter months immediately preceding and covering the epidemic wave of plague, and the other a smaller one during the rainy season, whereas *X. astia* has only one main wave of reproduction in the rainy season extending to autumn. During this reproductive period *X. cheopis* females increase greatly in proportion than the males but the condition is reversed in case of *X. astia*. These findings have been fairly corroborated in the field as well. The propensity of biting is however comparatively poorer in *X. cheopis* than in *X. astia* but it increases during the winter and spring.

As regards behaviour of rodents in the two wards (Ward 8, infected and Ward 10, uninfected) in Calcutta Ward 8 was found to be highly infected showing higher cheopis index and total flea population than in Ward 10. One of the reasons was the character of the housing conditions. The presence of bustees, godowns and huts in Ward 8 facilitated rat harbourage and infestations with house rats (*R. rattus*) more than in Ward 10. There were also large number of rat holes. On examination of these holes free fleas were found in 7.4% of the holes in April and in 18.5% in May, 21.0% in June and 0.0 in July and August. The specific flea indices were higher in Ward 8 than in Ward 10.

The result of resistance test also showed that all species of rats in Ward 10 were susceptible as against 60 per cent mortality in *R. rattus* examined towards the tail-end of the epidemic (1953-54) as shown in table 12.

**Table 12**  
*Resistance of rats caught in Wards 8 and 10 (1953-54);  
 Test dose-5000 Y. pestis*

| Type of rat          | Ward 8    |           | Ward 10   |             |
|----------------------|-----------|-----------|-----------|-------------|
|                      | No tested | No. died  | No tested | No died     |
| <i>R rattus</i>      | 30        | 18 (60)   | 22        | 22 (100 0)  |
| <i>R norvegicus</i>  | 28        | 26 (92 8) | 18        | 18 (100 0)  |
| <i>B bengalensis</i> | 50        | 49 (98 0) | 55        | 55 (100 0)  |
| Total                | 108       | 93 (86 0) | 95        | 95 (100 00) |

The rodent densities also differed between the two Wards as will be seen in table 13. The difference is obvious.

**Table 13**  
*Rodent densities in Wards 8 and 10 by species and season (1952-53)*

| Season      | Rodent densities in Ward 8 |      |     |      | Rodent densities in Ward 10 |     |      |      |
|-------------|----------------------------|------|-----|------|-----------------------------|-----|------|------|
|             | (endemic)                  |      |     |      | (non-endemic)               |     |      |      |
|             | Total                      | Rr   | Rn  | Bb   | Total                       | Rr  | Rn   | Bb   |
| Jan-March   | 22 6                       | 7 1  | 1 4 | 13 3 | 23 2                        | 8 6 | 0 7  | 13 9 |
| Apr-June    | 42 5                       | 18 1 | 6.1 | 10 2 | 10 2                        | 0 0 | 3 4  | 6 8  |
| July-Sept   | 21 8                       | 4 1  | 5.3 | 10 9 | 10 9                        | 2 7 | 2 1  | 6 1  |
| Oct-Dec     | 30 7                       | 5 4  | 2 7 | 32 1 | 12 1                        | 2 6 | 2 3  | 27 2 |
| Whole year  | 29 8                       | 8.5  | 3 9 | 22 5 | 12 5                        | 3 8 | 11.9 | 16 8 |
| Rats caught | 498                        | 148  | 65  | 291  | 445                         | 75  | 38   | 332  |

*Rr, R rattus, Rn, Rattus norvegicus, Bb, B bengalensis*

The above findings explain the existence of plague and its seasonal outbreak in one ward compared with another without plague.

## CONTINUOUS NATURAL TRANSMISSION EXPERIMENT

A continuous natural transmission experiment was carried out for 4 years through fleas among the two commonly available species of rats in Calcutta namely, *R. rattus* (domestic) and *B. bengalensis* (peridomestic). It was organised on the basis of experimental epidemiology and from the results obtained the following hypotheses were drawn regarding the mechanism of perpetuation of plague infection among the rodents bridging over the short and long intervals between epidemics (Seal 1961).

- (i) Apart from the clinical cases of plague among the rodent population infection can be maintained for prolonged periods in an inapparent or subclinical form in any of the partially immune rodents under investigation.
- (ii) Depending on the environmental and other ancillary conditions plague infection lying in the dormant state (usually in the spleen) rises as a relapse with bacteraemic condition facilitating transfer of infection through the ectoparasite, gradually leading to epizootic and mortality among rats. The infected fleas thus released then serve as the vector of plague infection in man.
- (iii) The fleas play only a secondary role in the matter of transmission. The infection is maintained actively in the rodents themselves.
- (iv) *X. cheopis* was the vector of choice for successful continuous transmission experiment while *X. astia* could propagate infection upto the third series.
- (v) In the perpetuation of infection organism of low grade virulence played more important role than the virulent organism. This virulence is usually regained by the circulation of the organism through susceptible hosts.
- (vi) The non-immunised rats exposed to infection but surviving for long periods retained the infection in larger numbers than the artificially immunised ones although the majority of either of them ultimately succumbed to the infection.
- (vii) Bacteraemia in the absence of impending clinical condition could not be established in these experiments as we were not sufficiently equipped to handle rats with bacteraemia before death.
- (viii) It was further observed that two doses of immunisation with specific plague protein Antigen A produced better immunity than

the killed Haffkine plague vaccine and *R. rattus* was found more vulnerable than *B. bengalensis*. Also, comparatively large number of non-immunised rats which survived for long periods were carrying *Y. pestis* in subclinical state than the immunised rats. The longest period observed was 180 days and the average 114 days. The final could not be reached due to the exhaustion of the experimental rats.

### FINAL LEAD TO ERADICATION

Having equipped with the above knowledge it became very easy to handle epidemic situation and also measures for eradication. The basic point is that plague being primarily a disease of the rats attempts should be intensively directed to eliminate infection from among the rats and accordingly the entire gamut of epidemiologic work was then directed to eradicate plague infection from amongst the rat population in the twin cities of Calcutta and Howrah.

The following steps were taken:

#### ***Control and Prevention of Plague*** (Seal 1980, 1987)

The various measures applied for the control of plague are:

- (i) The public were instructed how to handle dead rats
- (ii) Anti-rat measures against commensal rodents such as killing by mechanical means, and by use of predatory animals, trapping, fumigation, poisoning, baiting and other methods
- (iii) Rat-proofing of houses and godowns etc.
- (iv) Control of wild rodents
- (v) Vector control by DDT insufflation, application of other insecticides and fumigation
- (vi) Direct control of bubonic plague through hospitalisation of patients, management of contacts, evacuation, mass vaccination and chemo and sero-prophylaxis
- (vii) Control of pneumonic plague by case detection, isolation and treatment of patients, management of contacts, protection of staff and disinfection of contaminated objects
- (viii) Disposal of the dead
- (ix) Control of plague at a distance through quarantine laws and international intelligence system

- (x) The State Control Organization should search for the wild rodents and carry out rat destruction measures till the foci are completely eliminated.

*Precautions:* Vector control by pesticides by DDT or BHC spraying operations should be intensive and complete, because incomplete spraying often leads to resistance against the pesticides. The results of vector control operations in different parts of India has been published by Wagle and Seal (1953).

### CONTROL OF PNEUMONIC PLAGUE

Pneumonic plague is one of the easiest to control if precautionary measures are properly undertaken at the very onset.

- (i) Compulsory and complete isolation of active and suspected cases
- (ii) Cases should be attended under the cover of mask and by those who had received anti-plague vaccine
- (iii) Start intensive treatment with streptomycin, sulpha drugs and with highly potent anti-plague serum
- (iv) It may be supplemented by the usual anti-rat measures by cyanogassing and other measures
- (v) Protect others in the house and neighbourhood by administering anti-plague vaccine and chemoprophylaxis by sulpha drugs
- (vi) Provide all arrangements for laboratory diagnosis and special camp hospital
- (vii) Start strict surveillance operation.

### SURVEILLANCE OF PLAGUE

Surveillance is indispensable for areas where natural plague foci exist or where there is history of past plague infection. The method provides sound basis for plague control programme and indicate the need for long term ecological studies in crucial areas with a view to finally eradicate the disease. In fact, the surveillance operation undertaken by the author supported by the experimental work on rat movement, ecological studies of rats and fleas ultimately led to the eradication of the disease from India. The surveillance operation include the following steps (Seal & Modak 1975, 1976)

- (i) Trapping of rats from different areas systematically days after days and months after months and examination of samples of living and

dead rats for detection of plague organism and test them for virulence

- (ii) Identification of fleas, seasonal variation of density and breeding spree
- (iii) Insufflation of rat burrows and disinfestation of the premises of the affected wards and areas with DDT
- (iv) Regular and repeated cyanogassing of rat burrows
- (v) Detection of human cases, if any, their diagnosis and treatment and associated preventive and control measure
- (vi) Immunisation of persons living in the affected areas with anti-plague vaccine
- (vii) Carrying out ecological investigation and other researches on rodents fleas and improvement of bacteriological technique
- (viii) Carrying out health education measures particularly for eliciting cooperation of the people of the affected areas in handling dead rats and other methods of prevention

The above measures were continuously carried out through a plague control laboratory maintained for surveillance work for about 25 years till no trace of any infection among the trapped or dead rats collected daily in hundreds could be detected continuously at least for the last 10 years. In other words, the plague control laboratory was able to destroy several generations of rats in order to achieve this end. During this surveillance operation two new methods were discovered by the author namely (i) gel-precipitation test and (ii) new medium for sugar fermentation test (Seal 1972b). The gel containing specific soluble antigen and penicillin and streptomycin could identify the rats possessing immunity against plague infection indicating existence of infection among rat population. The advantage is that the test can be undertaken right in the field and hundreds of rats can be examined in a single day. Thus if the resistance test becomes continuously negative for 2 years it means absence of subclinical infection in the rat population or absence of plague infection (Seal 1972a).

#### ADDITIONAL MEASURES

During the period between 1959 and 1980 incidence of plague if reported anywhere in India was investigated under the guidance of the author e.g. in Mysore, Karnataka, Hyderabad and Assam (Gauhati) and eradication

measures undertaken. As a result no plague has been reported from any area in India from 1968 onwards.

### DIAGNOSIS AND TREATMENT OF PLAGUE (Seal 1980)

For arriving at a correct diagnosis and effective treatment the following steps are recommended

- (i) Collection of epidemiological data in a specially designed schedule by house to house visit
- (ii) Thorough clinical examination with particular attention to the sets of glands affected and their nature.
- (iii) Differential count of blood in certain cases and culture of blood and bubo materials
- (iv) Collection of sera from patients for serological examination
- (v) Animal inoculation test in a few instances only
- (vi) Collection of dead and live rats and fleas in the houses or nearby areas and examining them for physical evidence of plague by dissection and by examination of spleen smear and culture of splenic material and blood from heart. This would provide evidence of area infection and possible involvement of human beings
- (vii) Distribution of the classical features of four clinical varieties of the disease namely, (a) pestis minor, (b) bubonic, (c) septicaemic and (d) pneumonic to the field workers as well as to the hospital staff. Expecting pneumonic type might arise from the bubonic or septicaemic type the field workers should be instructed to look for the common features described in the text books.

Drugs which have been found effective for treatment are: (Seal 1960c)

|                                       | <i>Survival rate</i> |
|---------------------------------------|----------------------|
| Streptomycin antibiotic               | Not recorded         |
| Antiplague serum (Haffkine Institute) | 76.05                |
| Sulphapyridine                        | 73.0                 |
| Sulphathiazole                        | 77.1                 |
| Sulphathiazole and antiserum          | 80.0                 |
| Sulphamerazine                        | 97.5                 |
| Sulphadiazine                         | 91.2                 |

Appendix 1—The latest world position of reported incidence of plague cases between 1960 and 1982 is shown in figure 16.

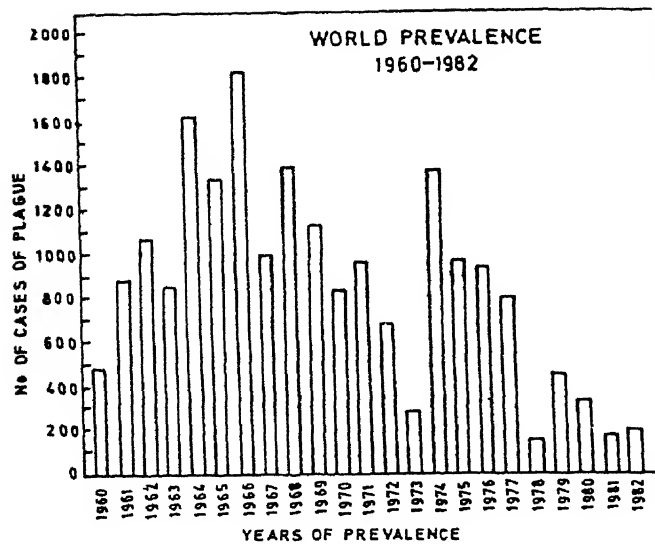


FIG 16

## References

1. Austrian Plague Commission 1895-1900 *Denk Schr Acad Wiss Wien* **66** Part I
2. Bacot A W and Martin C J 1914 *J Hyg (Camb)* **13** 209
3. Baker E E, Sommer H, Foster L E and Meyer K F 1947 *Proc Soc Exp Biol.* **64** 139
4. Bezsonova 1930 *Zbl Bakt (I Abt Orig)* **119** 32
5. Burrows T W, and Bacon C A 1956 *Brit J. Exp. Pathol* , **37** 461
- 6.. —, 1959 *Proc Diamond Jubilee Haffkine Institute, Bombay*, 14
7. Devignat R 1951 *W H O Bull.* **4** 247
8. —1958 *Rep presented before the W H O Expert Committee at Geneva*
9. Gauthier H C and Raybaud 1903 *Rev. Hyg Paris* **23** 426
10. German Plague Commission 1988 *Arab Gesundo Amt Berlin* 16
11. Indian Plague Res.Commission 1907 *J. Hyg. Camb.* 1907 Vols 1 to 12
12. Liston, W Glen 1906 *J Hyg. Camb.* **6** 425
13. Rao Raghavendra 1936 *Studies in Epidemiology of plague in Calcutta with special reference to long term periodicity.* DSc Thesis, Cal Univ.
14. Schutze H 1932 *Brit. J. Exp. pathol.* **13** 284
15. Seal S C 1942 *Ann. Rep. Haffkine Institute, Bombay*, 4
16. —1943 *Ann. Rep Haffkine Institute, Bombay*, p. 43



17. —and Stewart S E 1941 *Pub Hlth Rep.* **56** 792
18. —1950 *Ann Biochem Exp Med* **10** 99
19. —1951 a *Ann Biochem Exp Med* **11** 129, 143, 171
20. —1951 b *J Immunology* **67** 93
21. —1951 c *Proc Biol Med* **77** 675
22. —1952 *Ann Biochem Exp Med* **12** 123, 143
23. —1953 *J Immunol.* **71** 169
24. —1954 *Ann Biochem Exp Med* **14** 9
25. —1959 Envelope and Capsule of Plague bacillus. Paper presented before the 1st Microbiol. Congr held in Calcutta in October 1959
26. —1960a *W H O Bull* **23** 283, 293
27. —1960b *Indian J Med Sci* **14** 765
28. —1960c *Indian Med. Assoc* **35** 38
29. —1961 *Indian J Med Res* **49** 1019
30. —1972a *Indian J. Med Res* **60** 825
31. —1972b *Indian J Med Res* **60** 1414
32. —1976 *Indian J Pub Hlth.* **20** 25
33. —1980 D Sc thesis on Plague, University of Bombay
34. —1987 Plague conquest and eradication in India, ICMR Monograph 1987
35. —and Bhattacharjee L M 1961a *Indian J Med Res* **49** 974
36. —and—1961b *Indian J Med Res* **49** 1008
37. —and Modak K K 1975 *Indian J Pub Hlth.* **19** 126
38. —and — 1976 *Indian J. Pub Hlth.* **20** 25
39. —and Mukherje S P 1950 *Ann Biochem Exp Med* **10** 79
40. Sharif M 1951 *W H O Bull* **4** 75
41. Simond P L 1898 *Ann Instt Pasteur* **12** 625
42. Sokhey S S 1939 *Indian J Med Res.* **27** 313
43. —and Maurice H 1935 *Bull Off Int Hyg. Pub.* **27** 534
44. Wagle P M and Seal S C 1953 *W H O Bull* **9** 597
45. Wu-Lien-Teh 1926 A Treatise on Pneumonic Plague. (League of Nations Publication)
- et al. 1936 *Plague manual* (China: Shanghai)

**Indira Nath** (b. 14 January 1938) did M.D. (1969) from All India Institute of Medical Sciences, New Delhi; MRC (Path.) (1983), Royal College of Pathologists, UK. She is Professor and Head, Department of Biotechnology (1987- ) and was formerly Associate Professor (1986-87), Assistant Professor (1980-86); Department of Pathology, Lecturer, Department of Biochemistry (1972-80) at All India Institute of Medical Sciences, New Delhi.

Nath's main achievements are in cellular immunology with special reference to local problems. She and her coworkers evaluated immune responses in undernutritional states in experimental and human models, in individuals and primates immunized with antifertility vaccine involving HCG, and in infectious diseases such as tuberculosis and leprosy. They investigated immunoregulatory events in human leprosy using frontline technologies. Using both the natural *M. leprae* antigens and a unique recombinant antigen, LSR, which they identified, immune perturbation in leprosy patients was defined. They have identified B cell epitopes of LSR, which are uniquely recognized by the sera of ENL patients. She has contributed chapters to 7 books.

Nath is Fellow of National Academy of Sciences (India), Indian Academy of Sciences, Royal College of Pathologists (London), and National Academy of Medical Sciences (India); Life Member, Indian Association of Leprologists, Founder Life Member, Indian Immunology Society; Elected Fellow, College of Allergy and Applied Immunology; Member, Indian Association of Pathologists and Micro- biologists. She is the recipient of Dr Nitya Anand Endowment Lecture Award (INSA) (1987); Kshanika Oration Award (ICMR) (1984); ICMR-JALMA Trust Fund Oration Award (1981); Shanti Swarup Bhatnagar Prize (1983); Om Prakash Bhasin Foundation Award (1990).

*Indira Nath was elected to the fellowship of the Academy in 1992.*

**Address:** Professor & Head, Department of Biotechnology, All India Institute of Medical Sciences, New Delhi 110 029; Tel: (011) 6852286, 661123/453; Res: C3/10, Safdarjang Development Area, New Delhi 110 016; Tel: (011) 665977, 6851504.

## IMMUNOLOGICAL UNRESPONSIVENESS AND ITS REVERSAL IN LEPROMATOUS LEPROSY

INDIRA NATH

I am honoured to have been awarded the first Dr Nitya Anand Endowment Lecture by INSA. I first met Dr Nitya Anand when he chaired the discussions on the future strategies for research in Leprosy as part of the Eradication of Leprosy Programme initiated by Dr S Swaminathan Committee. My experiences during the meetings gave insight into the leadership quality of Dr Nitya Anand and his ability to encourage young persons like me to express our views

The studies I plan to describe synthesise some of our observations on the immunological features of leprosy as a natural response in man to exposure to the pathogen. They do not totally cover our studies of a decade and a half, but focus on these aspects which help us to draw some conclusions in continuum.

Leprosy is a unique infectious disease caused by the incultivable organism *Mycobacterium leprae*. Quarter of the world's population of leprosy patients live in India. The leprosy bacillus has a predilection for skin and nerves, resides within macrophages and Schwann cells and leads to a clinicopathological spectrum in man. Over the last decade and a half, it has been well established that the leprosy spectrum is due to the diverse nature of the host immune responses to *M. leprae*. At one pole of the spectrum lies tuberculoid leprosy (TT), which consists of a well circumscribed lesion, distinct from the surrounding healthy skin but affecting the neighbouring nerve. Such lesions show paucity of bacilli indicating thereby the ability of host to mount a healthy immune response capable of destroying the pathogen by granulomatous inflammation. In contrast, the other end of the pole is depicted by lepromatous leprosy (LL) wherein the patient shows disseminated multibacillary lesions scattered throughout the body and recognised microscopically by abundant collections of macrophages filled with leprae bacilli and lacking the accompanying lymphocytes seen in TT leprosy. Such individuals are difficult to cure and are a source of infection to the community. Between these two poles lie three clinical types of borderline leprosy which represent an unstable form of the disease. Depending on their clinical

proximity to the polar forms described above, they are variously designated as Borderline Tuberculoid (BT), Borderline Borderline (BB) and Borderline Lepromatous (BL) leprosy.

### IMMUNOLOGICAL FEATURES OF LEPROSY

It is well established that humoral immunity mediated by antibodies and cellular immunity affected by thymus-processed T-lymphocytes are the major limbs by which immunological reactions lead to protection or pathological reactions to foreign antigens. Since antibodies do not penetrate living cells, pathogens such as *M leprae* evade humoral immunity by their location within phagocytic cells. Tables 1 & 2 show the T and B cell functions in leprosy reported by us and others. TT patients have paucity of antibodies but mount good cellular immunity and are able to contain the spread of bacilli. LL patients have abundant general mycobacterial and specific anti *M leprae* antibodies and are yet unable to destroy the bacilli.

**Table 1**  
*Immunological features of Leprosy*

|                             | <i>Tuberculoid</i> | <i>Lepromatus</i> |
|-----------------------------|--------------------|-------------------|
| T Cells                     |                    |                   |
| Numbers                     |                    |                   |
| E rosettes                  | →                  |                   |
| CD 3+                       | →                  |                   |
| Functions                   |                    |                   |
| <i>Skin Tests</i>           |                    |                   |
| KLH, DNCB                   | →                  |                   |
| PPD                         | →                  |                   |
| <i>M. leprae</i>            | →                  |                   |
| <i>Lymphoproliferation.</i> |                    |                   |
| Mitogens                    | →                  |                   |
| Cross reacting ags          | →                  |                   |
| <i>M leprae</i>             | →                  |                   |
| <i>Leucocyte migration</i>  |                    |                   |
| <i>Inhibition</i>           |                    |                   |
| H 37 RA                     | →                  |                   |
| <i>M. leprae</i>            | →                  |                   |
| <i>IL 2 Production</i>      |                    |                   |
| Mitogens                    | →                  |                   |
| PPD                         | →                  |                   |
| <i>M. leprae</i>            | →                  |                   |

Antigen specific T cell anergy is longlasting. Depression in general T cell function is secondary and reversible

They lack cellular responses to *M. leprae* as indicated by unresponsiveness to the specific antigens both by *in vivo* skin tests and by *in vitro* lymphocyte function tests. Interestingly, these patients respond to other cross-reacting mycobacterial antigens. Indian LL patients during active disease show low general T cell functions in terms of T cell numbers and their responses to T cell mitogens. On treatment and decrease in bacillary load these defects are corrected (Nath et al. 1974, 1975).

**Table 2**  
*Immunological features of Leprosy*

|                                | <i>Tuberculoid</i> | <i>Lepromatous</i> |
|--------------------------------|--------------------|--------------------|
| HUMORAL                        |                    |                    |
| B cell nos                     | →                  | ↑                  |
| Serum Ig                       | →                  | ↑                  |
| <i>Mycobacterial abs</i>       |                    |                    |
| Other mycobacteria             | →                  | ↑                  |
| 65Kd                           | →                  | ↑                  |
| Arabinomannan                  | →                  | ↑                  |
| <i>M. leprae specific abs:</i> |                    |                    |
| FLAbs                          | →                  | ↑                  |
| Phenolic glycolipid            | →                  | ↑                  |
| MLO 4, 6                       | →                  | ↑                  |

Decrease with treatment and drop in bacillary load.

Secondary to disease.

Not protective in nature

May have a diagnostic role

That T cell mediated cellular immunity is of protective value has been clearly demonstrated in Leprosy. With a view to understanding the natural immune response to *M. leprae* our recent studies have centered around the characterisation of T cell lines and T cell clones in individuals who respond effectively to the leprosy bacillus. We have been able to enrich for leprae reactive T cells by developing long term T cell lines and cloning to homogeneity some of these T cells. All the clones we have developed to date bear the CD4<sub>+</sub> (helper/inducer) phenotype. Some of them produce the T cell growth factor (IL-2) and the lymphokine interferon gamma (INF-γ) thought to be an inducer of class II MHC antigens Ia and to have the ability to activate macrophages (Nathan et al. 1984). In terms of antigen specificity the T cell clones and lines fall into three groups: (i) crossreactive proliferative responses to 10-15 mycobacterial species and unrelated tetanus toxoid; (ii) with restricted

responses to 3-4 mycobacteria; and (iii) with responses to *M. leprae* and not to other mycobacteria. All the lines/clones required exogenous IL-2 for maintenance. Moreover, the time kinetics of IL-2 and INF- $\gamma$  varied not only within the same clone but also in response to various mycobacteria. We have not as yet produced T cell lines/clones from lepromatous patients.

### MECHANISMS UNDERLYING UNRESPONSIVENESS IN LEPROMATOUS LEPROSY

The predominant antigen specific anergy observed in LL patients is the most intriguing feature of leprosy. Whereas specific antibodies to the leprosy bacillus are abundant, the T cells are unable to respond to the same pathogen. The mechanisms underlying this defect have been difficult to detect since, ethical considerations limit the scope of investigation in man and there is no suitable animal model. Our studies have therefore been limited to the peripheral blood or skin lesions.

Anergy or lack of responsiveness to an antigen may be due to various factors, such as (i) genetic predisposition, (ii) active suppression due to cells or soluble factors, (iii) lack of antigen reactive cells due to clonal deletion, (iv) defects in antigen presentation, and (v) reduction in growth factors. We explored in-depth the role of active suppression and the growth factors in patients with leprosy.

### IMMUNOLOGICAL SUPPRESSION IN LEPROSY

#### *Suppressor T-Cells*

Subsets of T cells with phenotypic characteristics play an immunoregulatory role by exerting positive help to promote an immune response or by suppressing an ongoing response. In disease states both these functions may play a pathological role. The role if any of suppressor T cells in inhibiting the responses in lepromatous leprosy was investigated by us using various methodologies. Peripheral blood T cells from leprosy patients, were induced to generate suppression by both a general mitogen Con A and the specific antigen *M. leprae* (Nath et al. 1979, 1980). It was consistently shown that suppressor T cell activity though detectable in tuberculoid leprosy was absent or low in lepromatous leprosy. It appeared that during the natural course of the disease, exposure to *M. leprae* induced suppressor T cells in individuals with good T cell responses, as part of a general T cell reactivity. The lack of these cells as seen in LL

patients may explain the lack of control leading thereby to overproduction of antibodies and auto-antibodies. We further confirmed these results using HLA-D compatible siblings from the Wardha area which had been typified for HLA haplotypes by Drs N K Mehra, M C Vaidya from AIIMS and Drs J J van Rood and R R DeVries from Leiden. T cells from LL patients were mixed with PBMC from his genetically matched tuberculoid or normal siblings. Such cocultures were stimulated with antigen to study the effect on lymphoproliferation. These studies indicated unequivocally that LL patients did not have suppressive T cells in circulation (Nath et al. 1980). Recent studies from our laboratory showed that phenolic glycolipid I considered to be a unique antigen of *M. leprae* (Hunter et al. 1982) when incorporated in the form of liposomes generated general suppression of lymphoproliferation *in vitro* (Prasad et al. 1987). This suppression was observed in both tuberculoid and lepromatous patients. We concluded that suppression by T cells was not the major mechanism underlying anergy in Lepromatous leprosy. Furthermore, the proposed role of PGL-1 as a unique suppressor epitope (Mehra et al. 1984) appeared unlikely to be the central mechanism responsible for the unresponsiveness as it showed a general suppressor role and did not explain the differential features of the two poles of the leprosy spectrum.

### *Suppressor adherent cells*

Using HLA compatible siblings, we observed that monocyte-rich adherent cells of LL patients inhibited antigen-induced T cell proliferation of tuberculoid patients and healthy responder individuals (Nath *et al.* 1980). Such inhibition was mediated by soluble factors released *de novo* into the culture medium by monocytes of untreated LL and not by cells from tuberculoid patients. The *de novo* factors were observed to be decreased in treated patients, but could be induced on exposure of monocytes to antigen (Sathish et al. 1983). Interestingly, *M. leprae* induced maximal suppression in comparison to 5 other mycobacteria. Moreover, addition of such factors to lymphocytes stimulated with various mycobacteria also showed maximal inhibition of *M. leprae* related lymphoproliferation. Further characterisation on HPLC and by radioimmunoassay indicated that arachadonic acid metabolites, PGE<sub>2</sub>, leukotrienes and thromboxane were present in greater amounts in suppressor factors from LL individuals and in low or nondetectable amounts in healthy and tuberculoid individuals. Interestingly, these factors inhibited the production of T cell growth factor IL-2. (Nath *et al.* 1984). It had no effect on its utilisation. Though these factors are currently being further characterised, we are inclined to believe

that monocytes from lepromatous patients during contact with *M. leprae*, produce multiple factors which may play a role in suppressing T cell growth factors. We are currently investigating whether T cells play a role in influencing the monocytes in this function in order to explain the antigen induced suppression observed in leprosy.

### *The Status of T Cell Growth factors in Leprosy*

Concurrently with our studies on suppression, efforts were made to evaluate the defect if any in the production of T cell growth factors. It is now well established that monocytes/macrophages on contact with antigen release interleukin 1 (IL-1) which in turn helps the T cells to produce interleukin 2 (IL-2) which is required for the clonal proliferation of T cells. In our hands, both tuberculoid and lepromatous patients showed normal ability to produce IL-1 on stimulation with the general mitogen PMA and the specific antigen *M. leprae*. However, IL-2 production by T cells of LL patients was low to absent (Nath 1986). Though these individuals produced IL-2 to other antigens such as PPD or to T cell mitogens, they were unable to do so when stimulated with the specific antigen. It is possible from our earlier data on monocyte suppression that the peripheral blood cells being a mixture of monocytes, T cells and B cells would show defect in IL-2 production due to concurrent inhibition by monocyte factors.

### REVERSAL OF T CELL ANERGY IN LEPROMATOUS LEPROSY

Since the above studies implied that LL patients may have T cells which were suppressed by monocyte factors and which failed to clonally expand due to IL-2 defect, studies were undertaken to reverse the T cell anergy *in vitro* and *in vivo*.

### IN VITRO MODULATION OF T CELL RESPONSES

#### *Role of monocytes*

Adherent cells from tuberculoid or healthy responder individuals were cocultured with HLA-compatible nylon wool purified T cells from LL siblings in the presence of leprae antigens. Five such sibling pairs showed efficient lymphoproliferation indicating the presence of antigen-reactive T cells in these lepromatous patients (Nath *et al*).



### *Role of autologous dendritic cells*

Since lepromatous individuals possess suppressive monocytes and transfer of monocytes from tuberculoid patients would not be operationally feasible as a therapeutic measure, we explored the possibility of replacing monocytes with other accessory cells capable of presenting antigen to T cells (earlier studies had shown and it is well established that T cells only see antigen processed by accessory cells and in the context of MHC Class II antigens). Peripheral blood of man contain 1% of Dendritic Cells (DC) which are rich in MHC class II antigens and which can reconstitute lymphoproliferative responses in the absence of monocytes (Mittal & Nath 1987). DC were obtained from 15 LL patients and reconstituted in concentrations varying from 0.1 to 10% with purified autologous T cells and *M. leprae*. Interestingly, significant lymphoproliferation was observed in 9 patients. More importantly, 14/15 patients showed production of INF- $\gamma$ , a macrophage activating lymphokine (Mittal *et al.* 1988).

### *Role of IL-2*

Since our above studies had shown a defect in IL-2 production, we sought to reverse this by addition of exogenous IL-2 to *M. leprae* stimulated peripheral blood lymphocytes of 41 LL patients. IL-2 was obtained from 3 sources (i) mitogen induced JR-4 cells (ii) constitutively released IL2 from the gibbon cell line MLA; and (iii) recombinant IL-2 obtained commercially. In general 60-65% of LL individuals showed varying levels of improvement in antigen induced lymphoproliferation. (Nath *et al.* 1984).

These investigations provide proof that modulation of cell interactions and provision of exogenous T cell growth factors can reverse the well documented antigen specific T cell defect *in vitro*. It would appear that many LL patients possess antigen reactive T cells with IL-2 receptors which can be triggered to clonal expansion by IL-2. That DC constituted T cells could release INF- $\gamma$  would suggest that T cells capable of microbicidal activity are present in many LL patients.

## IN VIVO EMERGENCE OF ANTIGEN REACTIVE T CELLS

### *Leprosy reactions*

Some BL and LL patients suffer from acute episodic reactional states called erythema nodosum leprosum (ENL) which are characterised by dermal nodules and systemic manifestations of fever, neuritis and

arthralgia. We undertook investigation of such patients both for (i) T cell functions in peripheral blood, and (ii) phenotypic characteristics of cells in skin lesions.

Surprisingly, the hitherto anergic LL patients during ENL reactions showed enhanced T cell functions in terms of lymphoproliferation and lymphokine production (Laal *et al* 1985). Moreover, the usual lymphopenic lesions showed entrance of CD4<sub>+</sub> helper inducer T cells in significant numbers (Narayanan *et al.* 1984). Such cells were seen in the dermis and epidermis. That these cells were in a functional state was indicated by the appearance of Ia or MHC Class II antigens on the previously negative keratinocytes. Since antigen presentation to T cells is linked to the presence of MHC Class II antigens and they can be induced by INF- $\gamma$ , we conclude that CD4<sub>+</sub> T cells entering ENL lesions is indicative of release of INF- $\gamma$  in the local site (Thangaraj *et al.* 1988).

These studies confirm the natural emergence of antigen reactive T cells into the peripheral blood and lesions of lepromatous patients. Their presence becomes discernible during ENL reactions which we feel represent a state of T cell perturbation. The factors that trigger their emergence and traffic into lesions are not clear. That it is of importance to identify these events is evident from the fact that ENL lesions show fragmentation of bacilli and indicate that LL patients are capable of killing *M. leprae* given the right stimulus.

#### *In vivo Injection of PPD*

With a view to further understanding whether the lack of CD4<sub>+</sub> helper/inducer T cells in LL lesions were due to a general defect in emigration and accumulation or represented specific unresponsiveness, we generated tuberculin reactions by injecting PPD into lesions. In a collaborative study with Drs G Kaplan and Z A Cohn of the Rockefeller University, the epidermis and dermis were studied for phenotypic markers of the various cells using monoclonal antibodies. Similar to the ENL lesions described above we noted that CD4<sub>+</sub> cells entered LL lesions, keratinocytes expressed MHC Class II antigens and showed increased layers. These results indicated once more that LL lesions were permissive to the entry of helper/inducer T cells, and that these cells released lymphokines such as INF- $\gamma$  and epidermal growth factors (Kaplan *et al.* 1986, 1987).

Thus, both the *in vitro* and *in vivo* studies indicate that many of the hitherto anergic lepromatous patients would be amenable to immunological modulation whereby the antigen specific unresponsiveness may be reversed. One consistent feature that emerged was that approximately 1/3 of the individuals showed little responsiveness under similar conditions. Strategies to promote the total LL population to show improvement in immunological reactions leading to the killing of intracellular *M. leprae* are being currently explored by us.

Showing foresight and determination, India has introduced multidrug regimen on a large scale to control leprosy. Current reports indicate that though these drugs have a significant impact in clearing bacilli from the majority of multibacillary patients, yet a small foci of individuals continue to harbour *M. leprae* for long periods. Apart from individuals morbidity, such foci pose a threat of infection to the community at large. Immunological modulation may be required as an adjunct to chemotherapy in these patients. Moreover, immunotherapy may reduce the length of treatment required with drugs. We hope that studies such as above would contribute strategies for the better control of not only leprosy but other infectious diseases where the intracellular location of pathogens makes eradication difficult.

Thus a decade and a half of careful dissection of the leprosy engima has given us some answers but has not provided total consensus. This enticing disease becons us to explore further the mystery that is Leprosy.

### ACKNOWLEDGEMENTS

The author would like to record her gratitude to clinical collaborators without whose help, interest and discussions such studies would have been difficult to conduct.

Thanks are due to the Departments of Dermatology of All India Institute of Medical Sciences, Safdarjang Hospital, Ram Manohar Lohia Hospital; Lok Nayak J P Narayan Hospital of Delhi; to the Gandhi Memorial Leprosy Foundation, Wardha; Leprosy Mission, Shahdra; and S M S Medical College, Jaipur.

### References

1. Hunter S W, Fujiwara T and Brennan P J 1982 Structure and antigenicity of the major surface glycolipid antigens of *M. leprae*; *J Bio Chem* 257 1072

2. Kaplan G, Witmer M D, Nath I, Steinman R M, Laal S, Prasad H K, Sarno E N, Elvers U and Cohn Z A 1986 The influence of delayed immune reactions on human epidermal Keratinocytes *PNAS* 83 349  
 —, Nusrat A, Witmer M D, Nath I and Cohn Z A 1987 the distribution and turnover of Langerhans cells during delayed immune responses in human skin; *J Exp Med* 165 763
3. Laal S, Bhutani L K and Nath I 1985 Natural emergence of antigen reactive T cells in lepromatous leprosy patients during erythema nodosum leprosum; *Inf Imm* 50(3) 887
4. Mehra V et al. 1984 Lymphocyte suppression in leprosy induced by unique *M leprae* glycolipid, *Nature* (London) 308 194
5. Mittal A and Nath I 1987 I Human T cell proliferative responses to particulate microbial antigens are supported by populations enriched in dendritic cells, *Clin Exp Immunol* 69(3) 611
6. Mittal A, Mishra R S and Nath I 1988 Accessory cell heterogeneity in Leprosy-Dendritic cells and not monocytes support T cell responses (submitted for publication)
7. Narayanan R B, Laal S, Sharma A K, Bhutani L K and Nath I 1984 Differences in predominant T cell phenotypes and distribution pattern in reactional lesions of tuberculoid and lepromatous leprosy; *Clin Exp Immunol* 55 623
8. Nath I, Curtis J, Bhutani L K and Talwar G P 1974 Reduction of subpopulation of T lymphocytes in lepromatous leprosy; *Clin Exp Immunol* 18 81  
 —, Jayaraman T, Sathish M, Bhutani L K and Sharma A K 1984 Inhibition of Interleukin-2 production by adherent cell factors from lepromatous leprosy patients; *Clin Exp Immunol* 53(3) 531  
 —, Narayanan R B, Mehra N K, Sharma A K and Gupta M D 1979 Concanavalin A induced suppressor activity in human leprosy, *Journ Clin, Lab Immunol* 2(4) 319  
 —, Sathish M, Jayaraman T, Bhutani L K and Sharma A K 1984 Evidence for the presence of *M leprae* reactive T cells in patients with lepromatous leprosy, *Clin Exp Immunol* 58(3) 522  
 — *in vitro* reversal of T cell unresponsiveness in lepromatous leprosy, *Lepr Review* 57 Supp 207  
 — Singh R 1980 The suppressive effect of *M leprae* on the *in vitro* proliferative responses of lymphocytes from patients with leprosy, *Clin Exp Immunol* 41 406  
 —, van Rood J J, Mehra N K and Vaidya M C 1980 Natural suppressor cells in human leprosy: the role of HLA-D identical peripheral lymphocytes and macrophages in the *in vitro* modulation of lymphoproliferative responses; *Clin Exp Immunol* 42 203

9. Nathan C F, Prendergast T J, Wiebe M E et al. 1984 Activation of human macrophages; Comparison of other cytokines with interferon- $\gamma$  *J Exp Med* **160** 600
10. Prasad H K, Misra R S and Nath I 1987 Phenolic glycolipid-I of *Mycobacterium leprae* induces general suppression of Con A responses unrelated to leprosy type; *J Exp Med* **165** 239
11. Sathish M, Bhutani L K, Sharma A K and Nath I 1983 Monocyte derived soluble suppressor factor(s) in patients with lepromatous leprosy; *Infect Immun* **42** 890
12. Thangaraj H, Laal S, Thangaraj I and Nath I Epidermal changes in reactional leprosy — keratinocyte Ia expression as an indicator of cell mediated immune responses *Int J Leprosy*, (In Press Sept.)



**Chhitar Mal Gupta** (b. 1 September 1944) did Ph.D.(1969) from Agra University. He is Director, Institute of Microbial Technology, Chandigarh (1992- ); and was formerly Head, Membrane Biology Division, Central Drug Research Institute, Lucknow (1986-92).

Gupta delineated the mechanisms that generate and maintain the transbilayer phospholipid asymmetry in erythrocyte membrane; developed a novel class of liposomes which resist their lysis by blood components and survive for long in blood circulation; developed efficient and novel methods for successful treatment of aspergillosis and drug-resistant malaria; and demonstrated the role of 70 KDa heat shock proteins in stabilizing the membrane-associated cytoskeleton in erythrocytes under heat or pathologic stress.

Gupta is Fellow of Indian Academy of Sciences and National Academy of Sciences (India). He is the recipient of INSA Medal for Young Scientists (1974); Shanti Swarup Bhatnagar Prize (1985), Ranbaxy Research Foundation Award (1985); Dr Nitya Anand Endowment Lecture Award (INSA) (1989); S.S. Mishra Memorial Award (National Academy of Medical Sciences) (1982).

*Chhitar Mal Gupta was elected to the fellowship of the Academy in 1989*

**Address:** Director, Institute of Microbial Technology, Post Box 1304, Sector 39-A, Chandigarh 160 014; Tel. (0172) 44285, 21232 (Off); Telex: 0395-369 IMT-IN; Fax: 91-172- 44252; E-mail: cmg@imtech.ernet.in, Res. Director's Bungalow, IMTECH Housing, Sector 39-A, Chandigarh 161 014; Tel. (0172) 44030, 28050.

## **LIPOSOMES IN DRUG TARGETING**

C M GUPTA

It is my great pleasure and privilege to deliver Dr Nitya Anand Endowment Lecture this year. Indeed, I am greatly honoured on being elected by the Indian National Science Academy for this Lectureship.

Dr Nitya Anand, an internationally known Indian chemist, happens to be my most favourite teacher with whom I have been very closely associated for the last several years. I am deeply indebted to him not only for providing me with constant encouragement and whole-hearted support throughout my career but also for his generosity and liberal attitude. It was he who, about 20 years ago, could correctly identify my deep interest in Modern Biology, and extended his timely help in my smooth change-over from Synthetic Organic Chemistry to Membrane Biology. Also, he is the person on whose advice I initiated my work in 'liposomes - as new drug delivery systems' way back in 1979. In this lecture, I have, therefore, decided to present to you the progress that we have made on his suggested problem.

Liposomes are sphere-like structure which possess internal volume and are basically made-up of phospholipids. During their formation, they are capable of accomodating water soluble substances in their internal volume and water insoluble materials within their phospholipid membrane. Besides being used as model cells in basic studies on Biological Membranes and Immunology, liposomes have also been explored for their possible use as drug/enzyme carriers in Biology and Medicine. Although we have been interested for the last several years in both basic and applied aspects of liposomes, in this lecture, I shall restrict myself only to those studies which particularly deal with suitability of liposomes as drug carriers in therapy.

Liposomes have been widely considered useful as carriers for delivering drugs and enzymes to specific cells in biophase (Finkelstein & Weissmann 1981, Gregoriadis 1983, Schneider 1985). However, their application in therapy has been largely restricted due to the following problems:

(a) their blood-induced lysis in circulation, (b) their major uptake by liver and spleen, (c) lack of appropriate methods for targeting them to

specific cell *in vivo*, and (d) lack of methods which ensure delivery of liposomal contents to the target cells. Unless these problems are resolved, the use of liposomes as drug/enzyme carriers would remain limited in therapy.

The major emphasis of my laboratory during the last decade has been on designing of liposomes that are stable and longer living in blood circulation as well as on development of suitable methods for liposome-based targeting of drugs to erythrocytes and macrophages. These two cell type were chosen, as we were interested to ultimately use these drug targeting methods for homing of drugs to erythrocytes in malaria and to macrophages in macrophage-based infections like leishmaniasis, tuberculosis, etc.

### DESIGN OF BLOOD-STABLE AND LONGER LIVING LIPOSOMES

Liposomes from blood circulation are captured mainly by liver and spleen, where they are metabolised by a specific class of enzymes, called phospholipases. Although there are five types (*viz* A<sub>1</sub>, A<sub>2</sub>, B, C & D) of phospholipases present in living organisms, phospholipase A<sub>2</sub> is the major enzyme responsible for metabolism of phospholipids in mammals (Kates 1972, Bosch 1980). It was therefore argued that if the phospholipid component of liposomes could be rendered resistant selectively to phospholipase A<sub>2</sub>, the half-life of the resulting liposomes may be significantly enhanced due to their slower metabolism and consequently their slower uptake in the liver and spleen. The main consideration in bringing out the proposed structural modification in the phospholipids component was that the modification must not affect the bilayer forming properties of the resulting modified phospholipid and also the properties of liposomes formed therefrom.

To achieve the said goal, it was thought most appropriate to selectively alter the stereoelectronic characteristics of the C-2 ester bond in diacylglycerophospholipids, *e.g* phosphatidylcholines (PC). The change at this centre could provide simultaneously the following two advantages: (i) may render PC resistant to the action of phospholipase A<sub>2</sub> and (ii) may alter the binding ability of liposomes with blood components. The latter advantage may be derived from the fact that the C-2 ester region on PC aligns at the bilayer interface (Buldt et al. 1978, Gupta et al. 1979, Pearson & Pascher 1979) and therefore any structural change in this region may alter the binding ability of liposomes with plasma proteins, the main



blood components responsible for liposome lysis in blood (Krupp et al. 1976, Scherphof et al. 1978, Chobanian 1979). After giving serious consideration to the existing knowledge on the PC conformation in bilayers (Buldt et al. 1978, Gupta et al. 1979, Pearson & Pascher 1979), we decided to insert one NH residue between carbon atom of the C-2 ester group and its adjoining hydrocarbon chain.

The preferred conformation of PC in bilayers is such that its C-2 ester region is aligned at the bilayer interface (figure 1A). It was, therefore considered logical that introduction of one NH residue adjacent to the C-2 ester carbon atom would greatly help in stabilizing the preferred conformation of PC due to - (a) the hydrophilic nature of the NH residue, and (b) the partial double bond character of the N-C bond (figure 1B). Consequently the modified phospholipids (carbamyl PC) were expected to form readily the closed bilayers (or liposomes).

Carbamyl PC with varying fatty acyl chain composition were synthesized in high yields by carbamylation of the corresponding lyso-PC with appropriate alkyl isocyanates in the presence of a catalyst, 4-(N, N-dimethyl) amino pyridine (Gupta & Bali 1981). The modified phospholipids thus obtained were not hydrolysed by phospholipase A<sub>2</sub> from various sources but could easily be degraded by phospholipase C, and formed liposomes upon their dispersion in water (Gupta & Bali 1981). The size and permeability behaviour of these liposomes were similar to those of the natural PC (Gupta & Bali 1981). Also, the thermal phase transition temperatures of the carbamyl PC were similar to those observed for the corresponding natural PC (Curatolo et al. 1982, 1985). However, the transition enthalpy for the carbamyl PC were significantly higher than of the natural PC (Curatolo et al. 1985), suggesting that the low temperature state of the carbamyl PC bilayers is more ordered than that of the natural PC bilayers.

The carbamyl PC structure in bilayers was further investigated by means of fluorescence and NMR spectroscopy. Using 1,6-diphenyl-1,3,5-hexatriene as the fluorescent probe, the acyl chain packing order in the carbamyl PC bilayers was found to be higher than that observed with the natural PC bilayers (Bhakuni and Gupta 1989). The increased packing order appeared to result from the ordered structure of carbamyl PC bilayers, as suggested by <sup>1</sup>H and <sup>13</sup>C-NMR studies (Bhakuni and Gupta 1989). These studies showed that carbamyl PC head-group was more ordered than natural PC head-group, possibly due to intramolecular

hydrogen bonding between the NH and the phosphate residues (figure 2). Also, it was demonstrated that cholesterol interacted differently with carbamyl PC as compared to natural PC, and the  $\beta$ -face of this sterol preferentially oriented with the carbamyl PC C-2 acyl chain in bilayers (Bhakuni and Gupta 1989).

These findings clearly indicated that insertion of one NH residue between C-2 ester carbon and the adjoining alkyl chain in PC has, as expected, conferred on the resulting phospholipid not only the phospholipase A<sub>2</sub> resisting property but also an ability to form tightly packed bilayers. To examine whether these features of the modified PC would affect the liposome stability in blood, we studied the liposome interaction with serum *in vitro* and blood *in vivo*. Interaction studies with serum were considered desirable, as other investigators established that serum exerted greater lytic effect than blood on liposomes (Kirby et al. 1980a, b 1988). The liposome stability was ascertained by measuring both the blood (or serum) induced leakage of the entrapped solutes and the phospholipid transfer from liposomes to plasma (or serum) proteins.

The serum-induced leakage was determined by measuring efflux of water soluble solutes (*viz* 6-carboxyfluorescein & <sup>14</sup>C-glucose) from the liposomes in presence as well as in absence of serum. The lipid transfer to serum proteins was measured after labeling the phospholipid component with radioactivity. Both the serum-induced leakage and phospholipid transfer were considerably reduced by replacing natural PC with carbamyl PC in liposomes (Gupta et al. 1981), indicating that the modification introduced in the PC structure has indeed enhanced the liposome stability in serum. Based on this work it was proposed that an appropriate structural modification of the C-2 ester region in PC could, in general, render the resulting liposome comparatively resistant to lysis by the serum proteins (Gupta et al. 1981), which has later been confirmed by other investigators (Hermetter & Paltauf 1983).

To examine whether the enhanced liposome stability in serum would affect the liposome survival in blood circulation, both the blood-stability and survival times were measured in blood circulation of the injected animals. These studies, like serum interaction studies, also revealed that carbamyl PC liposomes resisted their lysis by the blood components (Bali et al. 1983). Also, these liposomes had longer survival times than the natural PC liposomes in the blood circulation (Bali et al. 1983). Apart from the increased survival, the liposome uptake by the liver was

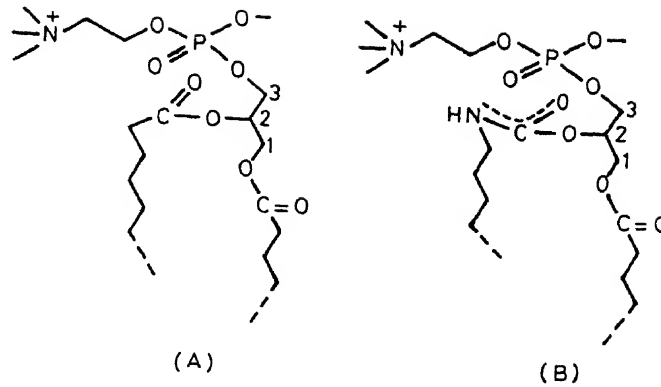


FIG 1 Conformation of PC in bilayers A. PC, B. Carbamyl PC

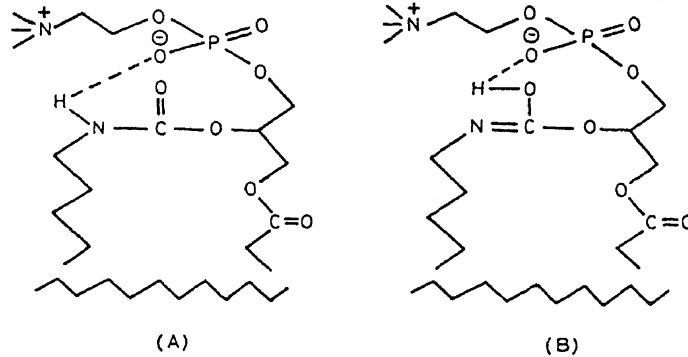


FIG 2 Schematic representation of intramolecular hydrogen-bonding possibilities in the carbamyl PC molecule.

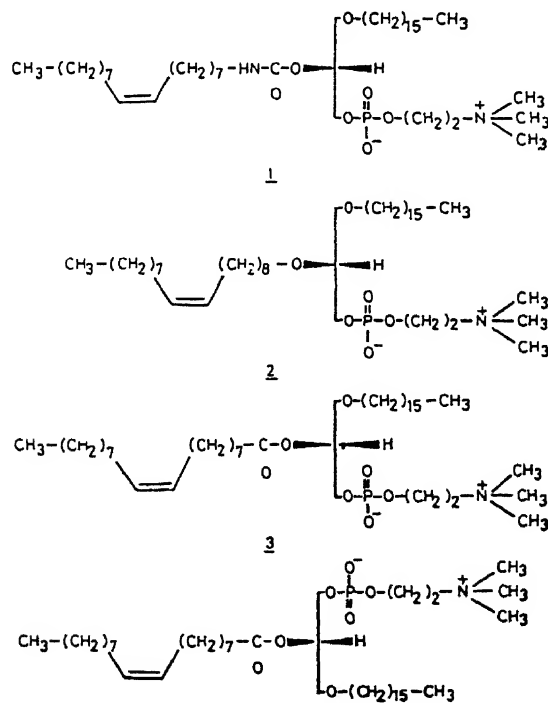


FIG 3 Carbamyl and ether analogs of PC (1-4)

decreased by replacing natural PC with carbamyl PC in liposomes (Bali et al. 1983, Gupta 1983).

The above results were consistent with our predictions made earlier based on the carbamyl PC structure. The increased stability of these liposomes may be considered to arise from the tighter phospholipid packing in the liposomes bilayer, which in turn could have resulted from the ordered carbamyl PC head-group structure. The latter structural feature of the phospholipid would not only make the bilayer interface structure highly ordered, but could also enhance the interlipid interactions. Consequently, the apolipoprotein, which is produced on collision of high-density lipoprotein with liposomes (Tall 1980), may not be able to insert its hydrophobic tail in the liposomes bilayer, thus leaving the liposomes largely undamaged. Likewise, the longer survival of these liposomes in blood circulation is perhaps related with the phospholipase A<sub>2</sub>-resisting property of the phospholipid component, which in turn could have affected the metabolism and consequently the capture of the liposomes by the liver.

Encouraged by these findings, we further synthesized (Agarwal et al. 1984) a few more modified phosphatidylcholines (figure 3) and tested the stability of their liposomes in serum *in vitro* (Agarwal et al. 1986a) and blood *in vivo* (Agarwal et al. 1986b). None of these modifications yielded liposomes better than the carbamyl PC liposomes.

Our findings summarised here strongly indicate that the liposome stability and survival time in blood circulation may be modulated by appropriately modifying the phospholipid structure. This is consistent with the earlier findings of other investigators who showed that replacement of PC with sphingomyelin in liposomes makes them more stable and longer living in blood circulation (Hwang et al. 1980, Allen 1981). Also the liposome stability in blood may be increased by increasing the cholesterol content of liposomes (Kirby et al. 1980, a, b, Tall 1980), suggesting that the tighter phospholipid packing may, in general, lead to the increased blood stability and longer circulation times of liposomes *in vivo*.

## DRUG TARGETING TO ERYTHROCYTES

As stated above, liposomes have inherent tendency to localize mainly in the reticuloendothelial (RE) cells. Therefore, for their homing to cells other than these cells, it is essential to graft on their surface the target cell-specific recognition markers, *e.g.* antibodies, lectins, sugars-etc. (Heath et al 1980, Leserman et al 1981, Ghosh et al 1982). By this technique a

significant, but not the total, fraction of liposomes may be expected to bind to the target cells. However, it is not necessary that all the cell-bound liposomes would successfully deliver their contents to the target cells, as the delivery depends on the mode of liposome interaction with cells (Pagano & Weinstein 1978).

There are several modes of liposome interaction with cells, but in the present context only the interactions that involve endocytosis or membrane-membrane fusion may be considered useful, as by these modes liposomal contents could quantitatively be delivered to the target cells (Pagano & Weinstein 1978). Since a mature mammalian erythrocyte is generally considered as a passive cell and possesses very little capacity to endocytose, we thought it of interest to examine whether it is possible to home the liposomal contents to these cells. Results of these experiments would enable us not only in assessing the suitability of liposome-mediated approach of drug targeting, but could also provide us with a method of drug delivery to pathologic erythrocytes, e.g. malaria parasite-infected cells.

Liposome binding to erythrocytes was increased by covalently attaching anti-erythrocyte antibody  $F_{(ab)2}$  fragments to the liposome surface. The antibodies to mouse (or rat) erythrocytes were raised in rabbits, and isolated from anti sera by affinity chromatography (Singhal et al 1986). The  $F(ab)_2$  fragments were covalently attached to the liposomes (Singhal et al 1986), and the antibody bearing liposomes thus obtained were used to evaluate their suitability as carriers in drug homing to erythrocytes.

The binding of liposomes to erythrocytes in whole blood (Singhal et al 1986) or *in vivo* (Singhal & Gupta 1986) was considerably (at least 20 fold) increased by covalently attaching anti-erythrocyte  $F(ab)_2$  to their surface. This binding depended not only on the liposome concentration but also on the antibody density on the liposome surface (Singhal et al 1986). Besides, the binding was fast (Singhal et al 1986, Singhal & Gupta 1986) and highly specific (Singhal et al 1986). Further, the binding process did not affect the liposome permeability, and also the survival of the liposome-bound erythrocytes in the blood circulation (Singhal & Gupta 1986). Moreover, a significant fraction of the cell-bound liposomes delivered their content to the target cells, presumably via membrane-membrane fusion (Singhal & Gupta 1986).

Only 10-15% of the injected dose of the antibody bearing liposomes interacted with the erythrocytes *in vivo*. Of these cell-bound liposomes, only 20-30% appeared to deliver their contents to the target cells (Singhal & Gupta 1986), which would amount to only 2.5% of the injected dose. It was therefore important to ascertain whether the amount of materials delivered to the cells by this method would be sufficient to control a disease. Keeping this in view, we entrapped the antimalarial drug chloroquine in the antibody targeted liposomes, and examined the efficacy of the liposomised drug against *Plasmodium berghei* infections in mice. Results of these experiments revealed that liposomised chloroquine was considerably more effective in treatment of the malarial infection as compared to the free drug (Agrawal et al 1987), demonstrating that inspite of the above limitations of the liposome-mediated drug delivery, the dose of drug delivered by this approach may still prove quite effective in combating the disease. In continuation of this work, we have now tested the effectivity of the delivered chloroquine in controlling chloroquine-resistant malarial infections. Results of these studies indicate that chloroquine delivered by this method could prove very effective in controlling the drug-resistant infections (*Unpublished results*).

All the above studies have been repeated and confirmed by the other investigators (Peeters et al 1988a, b, 1989), except for a small discrepancy between our (Singhal & Gupta 1986) and their (Peeters et al 1988a) results on kinetics of liposome clearance from the blood circulation. While we have reported longer survival times for our antibody bearing liposomes (Singhal & Gupta 1986), these workers have observed faster clearance of their liposomes from the circulation (Peeters et al. 1988a). This difference must mainly arise due to differences in the liposome preparations used in these experiments. We have used small unilamellar liposomes which are only slowly taken up by the RE system, whereas the other group prepared liposomes by the reverse phase evaporation method, which gives a mixture of large unilamellar and multilamellar liposomes. As the latter type of liposomes are preferentially and quickly cleared by the RE cells, the reported differences (Peeters et al 1988a) between our and their results are well within our expectations.

### DRUG TARGETING TO MACROPHAGES

Inspite of attaching cell-specific recognition markers to their surface, liposomes tend to localize mainly in the RE cells. It may therefore be expected that the infections that localize only in these cells could be

efficiently treated by administering the drug in liposomes. This efficiency must be further enhanced by grafting on the liposome surface the ligands that bind specifically to the RE cells. Keeping this in view, we designed and developed tuftsin bearing liposomes as carriers for drug homing to macrophages.

Tuftsin is a tetrapeptide (Thr-Lys-Pro-Arg) that resembles an integral component of IgG, and is released physiologically as the free peptide fragment after enzymatic cleavage (Nishioka et al 1972). This peptide is known to bind specifically to macrophages, monocytes and PMN leucocytes and also to potentiate the natural killer activity of these cells (Najjar 1987). Grafting of tuftsin on the liposome surface would therefore enable us not only in homing the liposomalised drug to macrophages but also to stimulate these cells nonspecifically against infections. Incorporation of tuftsin on the liposome surface was facilitated by attaching a fatty acyl residue to the C-terminus through an ethylenediamine spacer arm (Thr-Lys-Pro-Arg-NH  $(\text{CH}_2)_2$  -NH-CO-C<sub>15</sub>H<sub>31</sub>). The tuftsin bearing liposomes specifically recognised the target cells and delivered their contents to these cells (Singhal et al 1984). The liposome binding to cells was saturable and time dependent (Singhal et al 1984) and the cell-bound liposomes were apparently taken up by the cells presumably by the receptor-mediated endocytosis (Singhal et al 1984). Subsequent studies with the tuftsin bearing liposomes revealed that pretreatment of animals could render them resistant to malaria and leishmania infections (Gupta et al. 1986, Guru et al 1989). This resistance at least in leishmania infection appeared to arise from the macrophage activation, as the macrophages derived from the pretreated animals not only became refractory to infection but also failed to promote the growth and development of the intracellular parasite (Guru et al 1989).

Besides enhancing the nonspecific resistance against infections, the tuftsin bearing liposomes were found to be very useful also as drug carriers in leishmania therapy (Guru et al 1989). For these experiments sodium stibogluconate was encapsulated in both tuftsin bearing and tuftsin free liposomes, and efficacy of these drug preparations was evaluated against *Leishmania donovani* infections in hamsters. The drug delivered in tuftsin bearing liposomes was shown to be considerably more (at least 200 times) effective than the free drug in controlling the infection (Guru et al 1989).

These results thus clearly demonstrated that incorporation of tuftsin in the liposomes bilayer could render the liposomes very useful not only as drug carriers in leishmania therapy but also in enhancing the host's resistance to parasitic diseases. As leishmania parasite primarily resides within the mononuclear macrophages, the tuftsin bearing liposomes may prove very effective as vehicles for drug targeting also in other macrophage-based infections, like tuberculosis and leprosy.

### CONCLUDING REMARKS

In summary, we have established that an appropriate tailoring of the phospholipid structure can significantly increase the stability and survival times of the resulting liposomes in blood circulation. In addition, we have demonstrated that antibody targeted liposomes are capable of delivering drugs even to an inert cell, like the mature mammalian erythrocyte. Also, we have shown that incorporation of tuftsin in the liposomes bilayer renders the liposomes very useful not only as carriers for drug homing to macrophages but also for increasing the hosts's nonspecific resistance to parasitic infections.

Finally, it may be relevant to mention here that besides being useful in drug delivery, liposomes can also be used as vehicles for delivering protein antigens (Gregoriadis 1985) and genetic material (Fraley & Papahadjopoulos 1981) to the target cells. Liposome-associated cellular antigens not only exhibit better immunologic response than the free antigens but could also mimic the intact cell in terms of their biologic activity (Sengupta et al 1988, Narayana et al 1987).

### ACKNOWLEDGEMENTS

I must say that this lecture would have never been possible without the helps of my graduate students, Anu Bali (Puri), Kirti Agarwal, Arun Singhal, Vinod Bhakuni and Ajay Agarwal, who have made sustained efforts in making the programme successful. Also, I am grateful to my colleagues, Drs Anju Puri, Subhas Chandra and P Y Guru without whose expertise our working on malaria and leishmania could have been very difficult. Finally, I am very thankful to CSIR and ICMR for their financial support which made this work possible.



## References

1. Agarwal K, Bali A and Gupta CM 1984 Synthesis of carbamyl and ether analogs of phosphatidylcholines; *Chem. Phys. Lipids* **36** 169  
—, — and — 1986 a influence of phospholipid structure on the stability of liposomes in serum, *Biochim. Biophys. Acta* **856** 36  
, and Gupta C M 1986b Effect of phospholipid structure on stability and survival times of liposomes in circulation, *Biochim Biophys. Acta* **883** 468
2. Agrawal A K, Singhal A and Gupta C M 1987 Functional drug targeting to erythrocytes *in vivo* using antibody bearing liposomes as drug vehicles, *Biochem. Biophys Res Commun* **148** 357
3. Allen T M 1981 A study of phospholipid interactions between high-density lipoproteins and small unilamellar vesicles; *Biochim Biophys Acta* **640** 385
4. Bali A, Dhawan S and Gupta C M 1983 Stability of liposomes in circulation is markedly enhanced by structural modification of their phospholipid component; *FEBS Lett* **154** 373
5. Bhakuni V and Gupta C M 1989 Carbamylphosphatidylcholine cholesterol interactions in unilamellar vesicles; *Biochim Biophys. Acta* **982** 216
6. Bosch H V D 1980 Intracellular phospholipases A, *Biochim Biophys Acta* **604** 191
7. Buldt G, Gally H U, Seeling A, Seeling J and Zaccari G 1978 Neutron diffraction studies on selectively deuterated phospholipid bilayers, *Nature* **271** 182
8. Chobanian J V, Tall A R and Brecher P I 1979 Interaction between unilamellar egg yolk lecithin vesicles and human high density lipoprotein; *Biochemistry* **18** 180
9. Curatolo W, Bali A and Gupta C M 1982 Metastable phase behavior of a sphingolipid analogue; *Biochim Biophys. Acta* **690** 89  
—, — and — 1985 Phase behavior of carbamylphosphatidylcholine, a sphingolipid analogue; *J Pharm. Sci* **74** 1255
10. Finkelstein M C and Weissmann G 1981. Targeting of liposomes, in *Liposomes, From Physical Structure to Therapeutic Applications* ed. C G Knight p 443 (New York/North Holland: Elsevier)
11. Fraley R and Papahadjopoulos D 1981 New generation liposomes: The engineering of an efficient vehicle for intracellular delivery of nucleic acids; *Trends Biochem. Sci.* **6** 77
12. Ghosh P, Das P K and Bachhawat B K 1982 Targeting of liposomes towards different cell types of rat liver through the involvement of liposomal surface glycosides; *Arch. Biochem. Biophys.* **213** 266

13. Gregoriadis G 1983 Targeting of drugs with molecules, cells and liposomes; *Trends Pharmacol. Sci* **4** 304  
-- 1985 Liposomes for drugs and vaccines; *Trends Biotechnol* **3** 235
14. Gupta C M, Puri A, Jain R K, Bali A and Anand N 1986 Protection of mice against *Plasmodium berghei* infection by a tuftsin derivative, *FEBS Letts* **205** 351  
-- 1983 Modified liposomes: Their suitability as drug carriers; *Ann Natl Acad Med Sci (India)* **19** 22  
--, Bali A and Dhawan S 1981 Modification of phospholipid structure results in greater stability of liposomes in serum; *Biochim Biophys Acta* **648** 192  
-- and -- 1981 Carbamyl analogs of phosphatidylcholines: Synthesis, interaction with phospholipases and permeability behavior of their liposomes, *Biochim Biophys Acta* **663** 506  
-- Costello C E and Khorana H G 1979 Sites of intermolecular crosslinking of fatty acyl chains in phospholipids carrying photoactivable carbene precursors; *Proc Natl Acad Sci USA* **76** 3139
15. Guru P Y, Agrawal A K, Singha U K, Singhal A and Gupta C M 1989 Drug targeting in *Leishmania donovani* infections using tuftsin bearing liposomes as drug vehicles, *FEBS Letts* **245** 204
16. Heath T D, Fraley R T and Papahadjopoulos D 1980 Antibody targeting of liposomes: Cell specificity obtained by conjugation of F(ab)<sub>2</sub> to vesicle surface; *Science* **210** 539
17. Hermetter A and Paltauf F 1983 Interaction between ether glycerophospholipid vesicles and serum proteins *in vitro*, *Biochim Biophys Acta* **752** 444
18. Hwang K I, Luck K S and Beumier P L 1980 Hepatic uptake and degradation of unilamellar sphingomyelin/cholesterol liposomes: A kinetic study; *Proc Natl Acad Sci, USA* **77** 4030
19. Kates M 1972 Identification of individual lipids and lipid moieties, *Techniques in Lipidology* Chap 7, pp 565-569 (North-Holland: Elsevier)
20. Kirby C, Clarke J and Gregoriadis G 1980a Cholesterol content of small unilamellar liposomes controls phospholipid loss to high density lipoproteins in the presence of serum; *FEBS Letts* **111** 324  
--, -- and -- 1980b Effect of cholesterol content of small unilamellar liposomes on their stability *in vivo* and *in vitro*, *Biochem J* **186** 591
21. Krupp L, Chobanian A V and Brecher P I 1976 The *in vivo* transformation of phospholipid vesicles to a particle resembling HDL in the rat; *Biochem Biophys Res Commun* **72** 1251
22. Leserman L D, Machy P and Barbet J 1981 Cell-specific drug transfer from liposomes bearing monoclonal antibodies; *Nature* **293** 226

23. Najjar V A 1987 Biological effects of tuftsin and its analogs; *Drugs Future* 12 147
24. Narayanan R B, Ramu G, Sinha S, Sengupta U and Gupta C M 1987 Immunohistologic comparison between armadillo-derived leprosin and standard lepromin skin tests in leprosy patients; *Int Arch Allergy. Appl Immun* 82 202
25. Nishioka K, Constantopoulos A, Satoh P S and Najjar V A 1972 The characteristics, isolation and synthesis of the phagocytosis stimulating peptide tuftsin, *Biochem. Biophys Res Commun.* 47 172
26. Pagano R E and Weinstein J N 1978 Interactions of liposomes with mammalian cells, *Annu. Rev Biophys Bioeng* 7 435
27. Pearson R H and Pascher I 1979 The molecular structure of lecithin dihydrate; *Nature* 281 499
28. Peeters P A M, Oussoren C, Eling W M C and Crommelin D J A 1988a Immunospecific targeting of Immunoliposomes, P (ab')<sub>2</sub> and IgG to red blood cells *in vivo*, *Biochim Biophys, Acta* 943 137  
—, Claessons C A M, Eling W M C and Crommelin D J A 1988b Immunospecific targeting of liposomes to erythrocytes; *Biochem Pharmacol* 37 2215  
—, Brunink B G, Eling W M C and Crommelin D J A 1989 Therapeutic effect of chloroquine (CQ) containing immunoliposomes in rats infected with *Plasmodium berghei* parasitized mouse red blood cells Comparison with combinations of antibodies and CQ or liposomal CQ, *Biochim Biophys Acta* 981 269
29. Scherphof G, Roerdink F, Waite M and Parks J 1978 Disintegration of phosphatidylcholine liposomes in plasma as a result of interaction with high-density lipoproteins, *Biochim Biophys Acta* 542 296
30. Schneider M 1985 Liposomes as drug carriers-10 years of research, in *Drug Targeting* eds P Buri and A Gumma p 119 (Amsterdam' New York, Elsevier)
31. Sengupta U, Sinha S, Ramu G, Lavania R K and Gupta C M 1988 Soluble antigen of *M leprae* coupled with liposomes elicits both 'early' and 'late' delayed hypersensitivity skin reactions; *Int J Lepr* 56 45
32. Singhal A, Bali A and Gupta C M 1986 Antibody mediated targeting of liposomes to erythrocytes in whole blood; *Biochim. Biophys Acta* 880 72  
-- and Gupta C M 1986 Antibody mediated targeting of liposomes to red cells *in vivo*, *FEBS Letts* 201 321  
-- Bali A, Jain R K and Gupta C M 1984 Specific interactions of liposomes with PMN leukocytes upon incorporating tuftsin in their bilayers; *FEBS Letts* 178 109
33. Tall A R 1980 Studies on the transfer of phosphatidylcholine from unilamellar vesicles into high density lipoproteins in the rat; *J Lipid. Res.* 21 354

**Brahm Shankar Srivastava** (b. 1 June 1943) did Ph.D. He is Scientist and Head, Division of Microbial Genetics, Central Drug Research Institute, Lucknow; He was Formerly Professor & Chairman, Centre for Biotechnology, Jawaharlal Nehru University, New Delhi (1986-88).

Srivastava has worked with *Vibrio cholerae*, on its pathogenesis, antigens and cloning of antigenic genes leading to possibilities of Molecular vaccine. His findings are significant and quoted extensively. The novel observation on suppression of toxin biosynthesis by endogenous plasmids and demonstration of the role of plasmid-borne genes on attenuation of virulence was made by him. He revealed mechanism of adherence of vibrios to intestinal surface, its role in pathogenesis, identified and isolated the antigens, cloned its structural gene and demonstrated its role in protection as subunit vaccine. He showed RP4: mini-Mu replicon as vector for gene isolation and *in vivo* genetic engineering in *V. cholerae*.

Srivastava is the recipient of S.S. Bhatnagar Award in Medical Sciences (1984) and Dr. Nitya Anand Endowment Lecture Award (INSA) (1991).

**Address:** Scientist and Head, Division of Microbial Genetics, Central Drug Research Institute, Lucknow 226 001

## PATHOGENESIS OF CHOLERA AND VACCINE DEVELOPMENT

BRAHM S SRIVASTAVA

Among the most important enteric bacterial pathogens *Vibrio cholerae* 01 is responsible for epidemic and pandemic cholera in humans, particularly children living in poor sanitary conditions of underdeveloped countries. Infection occurs through oral route by ingestion of contaminated food and water. If vibrios successfully escape the acidic environment of the stomach, they arrive in the small intestine where a series of pathogenic events occur that result in colonization of vibrios and release of cholera toxin. The net result is the onset of diarrhea and associated clinical symptoms of cholera.

### PATHOGENESIS OF CHOLERA

Vibrios that survive the hostile environment of the stomach and arrive in the small bowel must be capable of multiplying in the lumen, increase in number and release toxin. They penetrate the mucous layer and come in close proximity to the epithelial surface of the intestine. *V. cholerae* possess a single polar flagellum which provides motile phenotype to this bacterium. Flagellum/motility is considered to be an important factor in the virulence of *V. cholerae* (Guentzel & Berry 1975). Nonmotile mutants neither penetrate the mucous layer nor exhibit virulence in experimental cholera (Schrack & Verwey 1976, Bhattacharjee & Srivastava 1979, Srivastava et al. 1980). Nonmotile aflagellate mutants poorly colonized the small intestinal epithelium. Inhibition of motility of vibrios by monoclonal and polyclonal antibodies *in vitro* and *in vivo* suggested that flagellar antigens might play a critical role in antibacterial immunity against cholera (Gustafsson & Holme 1985, Sinha et al. 1993).

Motile strains of *V. cholerae* appearing in large numbers in intervillous spaces and crypts adhere to the epithelial surface of the small intestine. The time course of adherence of vibrios to rabbit intestine and beginning of fluid outpouring in the ileal loop was studied (Nelson et al. 1976, Srivastava et al. 1980, Teppema et al. 1987). A good correlation was found between bacterial adherence and pathogenicity. Strains capable of adhering with high efficiency were pathogenic, whereas a nonadhesive

mutant isolated in our laboratory was found to be nonpathogenic (Srivastava & Srivastava 1980). Adherence of vibrios therefore to the mucosal cells of the intestinal epithelium is a very important step leading to colonization and release of toxin.

As the *V. cholerae* adhere and colonize the intestinal mucosa, they elaborate cholera enterotoxin which causes diarrhea (De 1959, Dutta et al. 1959). Cholera toxin is a protein whose structure, function and biological activity have been extensively studied. Briefly, it is composed of five B subunits and one A subunit. B subunit binds the toxin to the GM<sub>1</sub> ganglioside receptors on the mucosal enterocytes whereas the A subunit enters the enterocytes and activates irreversibly the adenylate cyclase system resulting in accumulation of cyclic AMP. Elevated level of cyclic AMP modifies the nature of cell membrane initiating fluid secretion from the cells (Cuatrecasas 1973, Gill 1976, see review Kaper & Srivastava 1992).

The role of cholera enterotoxin is undisputable in the pathogenesis of the organism. Other factors contributing to virulence were identified by testing a variety of mutants defective in motility, adherence and colonization (Bhattacharjee & Srivastava 1978, 1979, Levine et al. 1983). The studies suggested that immune mechanisms targeted forward inhibition of motility and adherence vibrios could be important in the prophylactic control of cholera.

Immunity to cholera involves antibodies directed towards vibrios (antibacterial) and the toxin released by them (antitoxin). Rabbits immunized with bacterial vaccine or toxin, or both, showed resistance to challenge. Antibacterial or antitoxin antibodies, or both, appeared in the circulation depending on the vaccines given to rabbits. There is convincing evidence that antibacterial and antitoxin immunities work synergistically (Svennerholm & Holmgren 1976, Srivastava et al. 1979). Antitoxin antibodies neutralize cholera toxin which appears to be the mechanism of antitoxin immunity (Peterson et al. 1979).

The data obtained in our laboratory strongly suggested that antibacterial immunity might act by inhibiting adherence of vibrios on the surface of the intestine (Srivastava et al. 1979, 1980, Jacob et al. 1993). A group of rabbits were immunized with killed whole cell bacterial vaccine to induce antibacterial immunity and were then challenged with a pathogenic *V. cholerae* strain in the ileal loop model (De & Chatterji 1953). When the negative and positive loops were examined for the

number of vibrios adherent to the intestine, it was found that few vibrios (2%) were adherent in the negative loops whereas a high percentage of vibrios were adherent in the positive loops. Thus there was a good correlation between low adherence and resistance to challenge. This finding suggested that antibacterial immunity provided protection mainly by interfering with the adherence of vibrios to intestine and that the adhesive antigen could be one of the key antigens in cholera immunity.

### ADHESIVE ANTIGENS

In our laboratory a nonadhesive mutant (CD11) of *V. cholerae* was isolated from its parent strain (KB207). Although CD11 was motile, chemotactic and toxinogenic, it adhered poorly, exhibited reduced virulence in experimental cholera and did not colonize the gut of infant mice (Srivastava & Srivastava 1980, Srivastava et al. 1980, Jacob et al. 1993). Analysis of the mutant revealed that a protein of 33 kDa present in KB207 was absent in CD11. Antibodies to this antigen inhibited adherence of KB207 to rabbit intestinal mucosa and colonization in an infant mouse model (Jacob et al. 1993). Our data suggested that the 33 kDa antigen could be an important antigen involved in adherence and colonization of vibrios in the intestine. A number of factors including fimbriae, hemagglutinins and cell envelop proteins have been reported to have possible relationship to *V. cholerae* adherence and colonization (Levine et al. 1983 for comprehensive review). Recently a pilus colonizing factor of 20.5 kDa from a classical strain of *V. cholerae* has been reported (Taylor et al. 1987).

### FLAGELLAR ANTIGENS

Little is known about the structure and composition of flagellum of *V. cholerae* and the antigens which are uniquely associated with the flagellum and likely to be exposed to the immune system of the host. In our laboratory, comparison of a nonflagellate nonmotile mutant with its parent strain revealed two proteins of 40 and 38 kDa associated with flagellum (Sinha et al. 1993). Antibodies to these antigens bind to flagella and inhibit motility of *V. cholerae*. Few other studies on the antigenic composition of flagellar sheath and core of *V. cholerae* are rather contradictory (Yang et al. 1977, Hranitzky et al. 1980, Richardson & Parker 1985).

## CHOLERA VACCINES

Soon after the identification by Koch in 1883 of the causative bacterium of cholera, attempts began to prepare vaccines to prevent cholera which is still continuing. As a result a series of immunizing agents have been produced and each vaccine suffered serious drawbacks. Some of these vaccines are (1) killed whole-cell vaccines, (2) toxoid vaccine, (3) combined vaccines and (4) attenuated *V. cholerae* vaccines. A comprehensive review on cholera vaccines is recommended (Levine et al. 1983).

Killed whole *V. cholerae* organisms have been used as parenteral as well as oral vaccines. There is some evidence to suggest that these vaccines are effective for short period against challenge with pathogenic *V. cholerae* and they seem to stimulate serum vibriocidal antibodies. Toxoid vaccines stimulate antitoxin immunity and include formadehyde or glutaraldehyde treated cholera toxin, purified B subunit and procholeraegenoid. Detoxified cholera toxins were not acceptable because of reversion to toxicity and poor protection. B subunit has been given parenterally and orally to human. It was found to be safe and stimulated antitoxin responses. In addition, these immunizing agents were tested in combination. The most significant study is the three oral doses of B subunit/killed whole cell vaccine given to volunteers participating in a vaccine efficacy challenge study. The results of these studies demonstrated so far complete safety and measurable degree of efficacy (Clemens et al. 1988). A number of naturally occurring attenuated *V. cholerae* 01 strains isolated from environmental sources in India and Brazil have been evaluated as vaccines with disappointing results (Cash et al. 1974). Several chemically mutagenized attenuated strains were isolated and one of them (Texas Star-SR) has been extensively studied in volunteers (Honda & Fuikelstein 1979). This vaccine too has several unacceptability problems. A genetically engineered strain of *V. cholerae* (Kaper et al. 1984) is under extensive field evaluation.

In our laboratory, two classes of attenuated strains of *V. cholerae* were isolated by two different techniques. One was obtained by conjugation between a wild type donor strain of *V. cholerae* and a slow growing, less adhesive and attenuated mutant strain. The attenuated mutant strain was isolated by mutagenic treatment and served as recipient in bacterial cross. Among the recombinants, nontoxinogenic attenuated recombinants were selected. These recombinant strains were avirulent in



different animal models of cholera. They were motile, adhered to intestine and colonized the gut of infant mouse and provided protection to challenge of pathogenic *V. cholerae* in rabbit ileal loop model. Because of these characteristics, the vaccine strain designated as *V. cholerae* CD1 and CD3 were proposed as candidate vaccine strains (Srivastava et al. 1979).

The other class of attenuated vaccine strain developed in our laboratory was obtained by transfer of P and V plasmids into pathogenic strains of *V. cholerae* (Sinha & Srivastava 1979). We made a novel observation and developed a novel technique of converting a virulent strain into an avirulent strain. This was achieved by introducing P and V plasmids into virulent strains (Sinha & Srivastava 1978a). It was shown that cells harbouring these plasmids became attenuated and did not cause experimental cholera. The loss of virulence was found to be due to decreased or suppressed toxin production (Khan et al. 1985). Such attenuated strains appeared attractive candidate vaccine strains if it could be shown that the plasmids are stable and immunogenicity of plasmid harbouring strains was not altered. The stability of the plasmid was confirmed *in vitro* and *in vivo* and excellent protection from challenge was observed in experimental cholera (Sinha & Srivastava 1978b, 1979).

A new class of cholera vaccine may be developed which will be composed of protective surface antigens of *V. cholerae* capable of generating a strong antibacterial immune response. The 33 kDa protein, identified as adhesive and colonizing antigen, was isolated and tested as vaccine in rabbits. It has been found that 33 kDa protein was immunogenic and conferred protection in rabbits and in combination with B subunit of cholera toxin, full protection was observed (Jacob et al. 1993). Thus 33 kDa protein has demonstrated its potential for use as a subunit vaccine against cholera in combination with B subunit. Such a vaccine is expected to be most desirable, as it would induce antibacterial as well as antitoxin immunities without fear of reversion to toxicity and toxic factors which are associated with live vaccine strains. It has already been shown that *V. cholerae* express more than one type of toxin, haemolysins and the cholera toxin gene could be amplified and more than one copy of the gene could be present in the genome (for review, Kaper & Srivastava 1992).

Alternatively structural genes encoding protective antigens could be cloned and expressed in heterologous systems (Pearson & Mekalanos 1982. Srivastava et al. 1985).

## ACKNOWLEDGEMENTS

I am grateful to my colleague for their collaboration and the Director for facilities. I wish to dedicate this manuscript to Dr Nitya Anand whose association since 1976 was most scientifically rewarding.

## References

1. Bhattacharjee J W and Srivastava B S 1978 Mannose-sensitive haemagglutinin in adherence of *Vibrio cholerae* eltor to intestine; *J Gen Microbiol* **107** 407-410  
— and — 1979 Adherence of the wild-type and mutant strains of *Vibrio cholerae* to normal and immune intestinal tissue; *Bull W H O.* **57** 123-128
2. Cash R A, Music S I, Libonati J P, Schwartz A R and Hornick R B 1974 Live oral cholera vaccine : evaluation of the clinical effectiveness of two strains in humans, *Infect Immun* **10** 762-764
3. Clemens J D, Harris, J R, Sack D A et al. 1988 Field trial of oral cholera vaccines in Bangladesh : results of one year of follow up; *J Infect Dis* **158** 60-69
4. Cuatrecasas P 1973 Gangliosides and membrane receptors for cholera toxin, *Biochemistry* **12** 3558-3566
5. De S N 1959 Enterotoxicity of bacteria-free culture filtrate of *Vibrio cholerae*, *Nature* **183** 1533-1534
6. De S N and Chatterjee D N 1953 An experimental study of the mechanism of action of *Vibrio cholerae* on the intestinal mucous membrane, *J Path Bacteriol* **66** 559-562
7. Dutta N K, Panse M V and Kulkarni D R 1959 Role of cholera toxin in experimental cholera; *J Bacteriol* **78** 594-595
8. Gill D M 1976 The arrangement of subunits in cholera toxin; *Biochemistry* **15** 1242-1246
9. Guentzel M N and Berry L J 1975 Motility as a virulence factor for *Vibrio cholerae*; *Infect Immun* **11** 90-97
10. Gustafsson B and Holm T 1985 Rapid detection of *Vibrio cholerae* O1 by motility inhibition and immunofluorescence with monoclonal antibodies; *Eur J Clin Microbiol* **4** 291-294
11. Hranitzky K W, Mulholland A, Larson A D, Eubanks E R and Hart L T 1980 Characterization of flagellar sheath protein of *Vibrio cholerae*; *Infect Immun* **27** 597-603
12. Honda T and Finkelstein R A 1979 Selection and characteristics of a *Vibrio cholerae* mutant lacking the A (ADP-ribosylating) portion of the cholera enterotoxin; *Proc Natl Acad Sci USA* **76** 2052-2056

13. Jacob A, Sinha V B, Sahib M K, Srivastava R, Kaper J B and Srivastava B S 1993 Identification of a 33 kDa antigen associated with an adhesive and colonizing strain of *Vibrio cholerae* El Tor and its role in protection; *Vaccine* **11** 376-382
14. Kaper J B, Lockman H, Baldini M M and Levine M M 1984 A recombinant nontoxinogenic *Vibrio cholerae* strains as attenuated cholera vaccine; *Nature* **308** 655-658  
— and Srivastava B S 1992 Genetics of cholera toxin; *Indian J Med Res A* **95** 163-168
15. Khan A A, Srivastava R, Sinha V B and Srivastava B S 1985 Regulation of toxin biosynthesis by plasmids in *Vibrio cholerae*; *J Gen Microbiol* **131** 2653-2657
16. Levine M M, Kaper J B, Black R E and Clements M L 1983 New knowledge on pathogenesis of bacterial enteric infections as applied to vaccine development; *Microbiol Rev* **47** 510-530
17. Nelson E T, Clements J D and Finkelstein R A 1976 *Vibrio cholerae* adherence and colonization in experimental cholera electron microscopic studies; *Infect Immun* **14** 527-547
18. Pearson G D N and Mekalanos J J 1982 Molecular cloning of *Vibrio cholerae* enterotoxin genes in E coli K-12; *Proc Natl Acad Sci USA* **79** 2976-2980
19. Peterson J W, Hejtmancik K E, Markel D E et al. 1979 Antigenic specificity of neutralizing antibody to cholera toxin; *Infect Immun* **24** 774-779
20. Richardson K and Parker C D 1985 Identification and occurrence of *Vibrio cholerae* flagellar core proteins in isolated outer membrane; *Infect Immun* **47** 674-679
21. Schrank G D and Verwey W F 1976 Distribution of cholera organisms in experimental *Vibrio cholerae* infections : proposal mechanisms of pathogenesis and antibacterial immunity; *Infect Immun* **13** 195-203
22. Sinha V B and Srivastava B S 1978a Suppression of pathogenicity by P and V plasmids in *Vibrio cholerae*; *J Gen Microbiol* **104** 251-255  
— and — 1978b Plasmid induced loss of virulence in *Vibrio cholerae*; *Nature* **276** 708-709  
— and — 1979 Attenuation of virulence by P and V plasmids in *Vibrio cholerae* strains suitable for oral immunization; *Bull W H O* **57** 643-647  
—, Jacob A, Srivastava R, Kaper J B and Srivastava B S 1993 Identification of the flagellar antigens of *Vibrio cholerae* El Tor and their role in protection; *Vaccine* **11** 372-375
23. Srivastava B S, Sinha V B and Srivastava R 1979 Attenuated recombinant strains of *Vibrio cholerae* for oral immunization; *Bull W H O* **57** 649-654

24. Srivastava R and Srivastava B S 1980 Isolation of a nonadhesive mutant of *Vibrio cholerae* and chromosomal localization of the gene controlling mannose-sensitive adherence; *J Gen Microbiol* 117 275-278  
—, Sinha V B and Srivastava B S 1979 Re-evaluation of antibacterial and antitoxin immunities in experimental cholera; *Indian J Med Res* 70 369-373  
—, —, and — 1980 Events in the pathogenesis of experimental cholera role of bacterial adherence and multiplication; *J Med Microbiol* 13 1-9  
—, Khan A A and — 1985 Immunological detection of cloned antigenic genes of *Vibrio cholerae* in *Escherichia coli*; *Gene* 40 267-272
25. Svennesholm A M and Holmgren J 1976 Synergistic protective effect in rabbits of immunization with *Vibrio cholerae* lipopolysaccharide and toxin/toxoid; *Infect Immun* 13 735-740
26. Taylor R K , Miller V L, Furlong D B and Mekalanos J J 1987 Use of phoA gene fusions to identify a pilus colonizing factor coordinately regulated with cholera toxin; *Proc Natl Acad Sci USA* 84 2833-2837
27. Teppema J S , Guinee P A M, Ibrahim A A, Paques M and Ruitenberg E J 1987 In vivo adherence and colonization of *Vibrio cholerae* strains that differ in hemagglutinating activity and motility, *Infect Immun* 55 2093-2097
28. Yang G C H, Schrank G D and Freeman B A 1977 Purification of flagellar cores of *Vibrio cholerae*; *J Bacteriol* 129 1121-1128



**Bimal Kumar Bachhawat** (b. 16 August 1925) did Ph.D. (1953) from University of Illinois, USA. He is Bhatnagar Fellow (1990- ) and was formerly Dean, (1988-90) at Faculty of Interdisciplinary and Applied Sciences, University of Delhi, South Campus, New Delhi.

Bachhawat's discovery of the biosynthetic pathway of cerebroside-3-sulphate led to the understanding of the mechanism of the synthesis of this complex glycolipid. The original finding that the arylsulphatase A is the deficient enzyme in the in-born error of metabolism (metachromatic

leucodystrophy) led to the development of a simple diagnostic method for this disease. His discovery (with M.J. Coon) of the enzyme HM6COA lyase led to the basic understanding of the ketone body synthesis in animals. The discovery of a new enzyme CMP-n, acetylneuraminic acid degrading enzyme, shed light on its regulatory role in the biosynthesis of membrane sialic acid. The receptor-ligand interaction with the liposome as a model membrane led to the understanding of the density of receptor on the cell surface and its role in ligand interaction. Recently, Bachhawat and his group have developed a liposomal formulation of Amphotericin-B which can effectively cure systemic aspergillosis in experimental animals. He has Authored more than 20 books/chapters.

Bachhawat is Fellow of Indian Academy of Sciences and National Academy of Sciences (India). He was Member, Council (1975-77) and Vice-President (1987-88), INSA. He is the recipient of Shanti Swarup Bhatnagar Prize (1962), Amrut Mody Research Award (1974), Golden Jubilee Gold Medal (Institute of Science) (1976); J.C. Bose Award (1980); Bashambar Nath Chopra Memorial Lecture Award (INSA) (1977); FICCI Award (1982); B C Guha Memorial Lecture Award (INSA) (1984); Birla Smarak Khosh (1986); Shanti Swarup Bhatnagar Medal (INSA) (1991); Padma Bhushan (1990).

*Bimal Kumar Bachhawat was elected to the fellowship of the Academy in 1973.*

**Address:** Bhatnagar Fellow, Department of Biochemistry, University of Delhi, South Campus, New Delhi 110 021, Tel: (011) 678876, 6882231 (Off); 606207 (Res); Fax 91-11- 6885270.

## **LIPOSOME TECHNOLOGY**

**B K BACHHAWAT FNA**

Mr President, Distinguished Fellows of the Academy, Ladies and Gentlemen!

It is indeed an honour and privilege for me to deliver the Professor Biresch Chandra Guha Memorial Lecture this year. I am grateful to the Indian National Science Academy for electing me for this Lectureship.

Professor Biresch Chandra Guha, the doyen of Indian biochemists, was my teacher and introduced me to the subject of biochemistry decades ago. Professor Guha was not only an ideal teacher and a renowned biochemist, he was also a futurist and a great seer. I remember as early as in the fifties he used to discuss with us about the potential use of enzymes in biological industries. At a time when the term biotechnology was unheard of and even modern biology was yet to penetrate into industrial sector, Professor Guha visualised the great potential of biotechnological revolution which is now knocking at our door these days. I imagine how happy he would have been to see his prophesy coming true had he been living today. I, therefore, embark upon the unique privilege of delivering this lecture with a deep sense of humility and gratitude.

Liposomes are used extensively in the studies on model membranes, as a drug and enzyme carrier and in immunology. In this lecture, I will discuss the work carried out in my laboratory in the following areas:-

- (A) Liposome as a model membrane.
- (B) Liposome as a drug and protein carrier.
- (C) Liposome as an adjuvant in Immunology
- (D) Liposome as a haptenic carrier for small molecules.

### **A. LIPOSOME AS A MODEL MEMBRANE**

Cell surface carbohydrates are implicated in a number of important cell surface properties such as cell-lectin, cell-toxin and cell-cell interaction. Since the intact cell is a very complicated system and as such the study on molecular level is very difficult, if not impossible. Accordingly, it is logical to study the individual component in a model membrane so that

one can at least get an idea of a particular component in relation to the cell-ligand interaction under defined experimental condition.

The receptor properties of a number of glycoproteins have been studied. However, the study of the receptor property of glycolipids is difficult since they form micelles in aqueous solution. This difficulty was successfully resolved by incorporating the glycolipid in the liposome. We developed a very simple system to study. In collaboration with a biophysicist, Dr S K Poddar of Bangalore and one of my brilliant graduate students, Dr A Surolia, we developed a simple system to study the lectin-glycolipid interaction.<sup>1-4</sup> Ganglioside (GM<sub>1</sub>) containing multilamellar and unilamellar liposomes were prepared and their interactions with galactose binding *Ricinus communis* lectin (RCA<sub>1</sub>), was investigated. Measurement of various kinetic parameters e.g. association constant and rate of cluster formation led us to conclude that the system could be used as a simple model for the study of receptor-ligand interaction. A few interesting inferences were derived from this study; (1) About 60 per cent of the galactose residues of GM<sub>1</sub> is externally available on the surface of unilamellar liposomes; (2) in case of multilamellar liposomes only 25 per cent of the same is exposed on the surface; (3) the rate of lectin GM<sub>1</sub> liposome interaction is markedly affected by the surface density of the sugar residue (number receptor sites per  $\mu\text{m}^2/\text{liposomes}$ ) i.e. by GM<sub>1</sub> concentration in liposomes. Thus the rate of liposome aggregation increased 20-fold when the molar ration of GM<sub>1</sub> to phospholipid was increased from 0.08 to 0.18. In another study, fatty acid chain length as well as oligosaccharide chain length on the rate of interaction between RCA, and glycolipid liposomes were investigated. These studies suggest that the phase transition temperature of the phospholipid component of liposomes, length of the surface-bound oligosaccharide chain and cholesterol concentration also affect the binding of the terminal sugar with the lectin.

These studies led us to conclude that (1) the terminal galactose residue is almost embedded into the lipid of liposomes containing Gal-Cer (2) lectin binding with Gal-cer and cytolipon H liposomes are strongly influenced by fatty acid chain length of the phospholipid compound as well as the cholesterol concentration of liposomes. The role of cholesterol was not clearly understood. It may be that cholesterol may affect the membrane fluidity. It was also noted that the increased receptor density can be predicted by the increased lectin cross binding. It may be of interest to mention here that although the lectin mediated agglutination of

liposome was very marked with density of the receptor, the binding of the lectin linearly increased with the increased concentration of the receptor on the surface. The impact of our initial report on the glycolipid lectin interaction can be gauged from the following quotation from a recent review of the subject by Grant and Peters<sup>5</sup>: "*Surolia, Bachhawat and Poddar have been responsible for a series of very clever experiments which set the pace.*"

## B. LIPOSOME AS A DRUG DELIVERY SYSTEM

In a collaborative work with Dr James H Austin of Denver Colorado, USA and my group at that time at Vellore, we for the first time showed that the deficiency of arylsulphatase was the main factor for the accumulation of cerebroside-3 sulphate in a sphingolipidosis, metachromatic leukodystrophy<sup>6,7</sup> Arylsulphatase A, a lysosomal enzyme was found to be a glycoprotein.<sup>8</sup>

In 1971, Raman and her coworkers initiated a number of studies where they used liposome encapsulated enzyme for the delivery into the cell.<sup>9</sup> During this period in a series of brilliant experiments, Ashwell as well as Morrell<sup>10</sup> and their groups in USA reported their work on the *in vivo* survival of glycoprotein and its removal through  $\beta$ -galactosyl moiety of the oligosaccharide chain of the glycoprotein by the liver. A specific receptor for  $\beta$ -galactoside moiety on the hepatocyte surface was recognised. A mannoside specific receptor was also recognised by Stahl *et al.* in 1978<sup>11</sup> on the cell surface of reticuloendothelial system of rats including the liver sinusoidal cells and macrophages.

These facts led me<sup>12</sup> to suggest that by modification of the liposomal surface by incorporating various glycosides it will be possible to target different tissues *in vitro*. With this aim in view we were able to construct a liposome model from our work on the kinetics of lectin-glycolipid interaction. This model shows that a part of the oligosaccharide chain is on the surface of the liposome and is available for the binding with lectins (Fig. 1).

With this background information, we initiated our *in vivo* experiments with GM<sub>1</sub> ganglioside and asialo GM<sub>1</sub>, ganglioside containing liposome with entrapped enzymes, proteins and drugs. Initially, we used a heterogeneous population of liposomes and we observed that the liposome-entrapped materials were very rapidly taken up by the liver. Using appropriate oligosaccharide residues as *in vivo* inhibitors, we



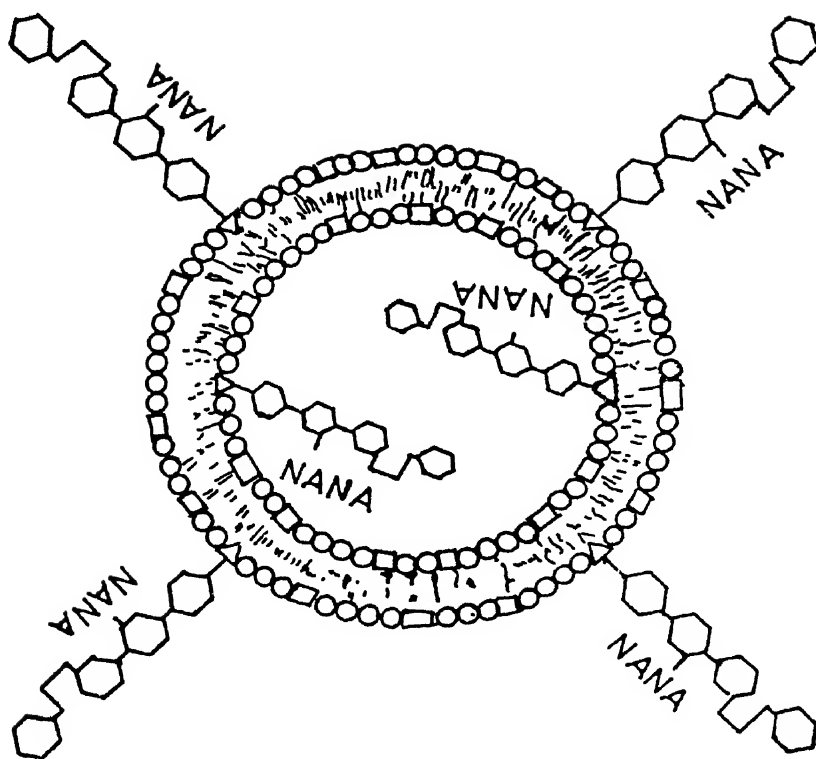


FIG 1 Schematic representation GM<sub>1</sub> containing single compartment liposome

concluded that the uptake by the liver was through sugar specific receptors<sup>13,11</sup> In an attempt to identify the types of the liver cells involved in this sugar-specific uptake, we administered *in vivo* asialo GM<sub>1</sub> liposomes and mannosylated liposome entrapped labelled  $\gamma$ -globulin. Livers were perfused and the different cells were isolated. It was observed that the galactosylated liposome entrapped material were enriched in the hepatocyte. This sugar specific enrichment was completely abolished<sup>15</sup> if asialofetuin was administered along with the asialo GM<sub>1</sub> liposomes. Similarly, mannosylated liposome entrapped materials were enriched in the Kupffer and endothelial cells and there was partial decrease in the uptake of the liposomes when mannan was a competitive inhibitor.

Using homogeneous small unilamellar liposomes (SUV), we carried out a series of experiments and developed a new method for the preparation of synthetic glycolipids<sup>16</sup>. In this method phosphatidyl ethanolamine was coupled with mellibiose or lactose through cyanoborohydride to form a synthetic glycolipid having terminal  $\alpha$ -galactosyl and  $\beta$ -galactosyl residue. It was observed that  $\beta$ -galactosylated liposomes are preferentially taken up by the liver. During this period a

number of workers, Gregoriadis,<sup>17</sup> Rahman and others<sup>18</sup> had attempted unsuccessfully to target the liposome having the glycoside residues. We suggested that their inability to target the liposome having the glycoside residues. We suggested that their inability to target the liposome might have been due to the low level of glycolipids they had incorporated in their liposome as it was observed by us that the sugar specific uptake of the liposome was dependent on the density of the sugar residue on the surface of the liposome<sup>19</sup>.

However, without taking into account that we had reported the liposomal uptake using heterogeneous preparation as well in other cases homogeneous small unilamellar liposomes, a number of reports<sup>20,21</sup> appeared in the literature which basically confirmed our observation that  $\beta$ -galactoside liposomes are specifically taken up by the hepatocytes. These investigators had used ceramide disaccharide having terminal  $\beta$ -galactoside. However, unfortunately since they have used different methods of preparation of liposomes and have not measured the available number of  $\beta$ -galactoside residues on the surface of these liposomes, the discrepancy regarding the rate of uptake of the liposomal entrapped material compared to our results would not be resolved.

A number of exciting probabilities of the use of  $\beta$ -galactosylated liposomes are now becoming apparent. We have been able to show that asialo GM<sub>1</sub> liposome entrapped material could be effectively used to prevent the hepatitis-like effect of D-galactosamine<sup>22</sup>. It is known that D-galactosamine can specifically destroy hepatocytes. This effect of D-galactosamine can be reduced or prevented by the administration of uridine. Since asialo GM<sub>1</sub> liposome entrapped uridine administered just prior to the administration of galactosamine could effectively prevent the galactosamine toxicity. Asialo GM<sub>1</sub> entrapped uridine was almost 10 times more effective compared to free uridine in combating the D-galactosamine toxicity.<sup>23,24,25</sup>

It may be mentioned here that asialo GM<sub>1</sub> liposome entrapped uridine was effective only when hepatocytes were intact i.e. prior to the administration of D-galactosamine and it was found to be comparatively ineffective when the hepatocytes were destroyed i.e. three hours after D-galactosamine administration. This may be due to the fact that D-galactosamine affects the hepatocyte leading to the probable loss of surface receptor for  $\beta$ -galactoside residue resulting in impaired receptor mediated uptake of the liposome.

Another interesting development of the application of  $\beta$ -galactosylated liposome had been in the direction of the introduction of genetic material into hepatocyte<sup>26</sup>. Scherph of and Nicolou jointly using ceramide lactoside were able to introduce prepronisulin gene was administered *in vitro* it generally accumulated into Kupffer cells in a degraded form. As has been pointed out by these investigators this is an interesting possibility as the hepatocytes are secretory cells.

In view of the above exciting possibilities of  $\beta$ -galactosylated liposomes, it was felt that further investigation of the nature of uptake of the liposome was important. Accordingly, the intracellular localization of liposome entrapped material both in the parenchymal and non-parenchymal cells of the liver was examined<sup>27</sup>. After taking proper precaution during the isolation of the liver; cells to minimize the receptor mediated uptake and lysosomal degradation of the liposomes during the liver perfusion it was observed that the enhanced uptake of asialo GM<sub>1</sub> liposome by parenchymal cells was due to the  $\beta$ -galactosylated liposomal cells. Sub-cellular studies indicated substantial lysosomal localization of the liposome entrapped material both in parenchymal and non-parenchymal cells. Asialofetuin inhibited specifically the uptake of asialo GM<sub>1</sub> liposome. Negatively charged liposome had also the enhanced uptake by the liver cells compared to neutral liposome but lower than galactosylated liposome. In the case of negatively charged liposome both parenchymal and non-parenchymal cells showed increased uptake. It was of interest that although there was inhibition of uptake of asialo GM<sub>1</sub> liposomes are cointernalized with asialofetuin through the common lysosomal route of ligand internalization or the desialated glycoprotein in the lysosome may somehow protect the liposome entrapped material from the highly metabolic milieu of lysosome. The exact mechanism of this phenomenon is yet to be known.

As mentioned earlier, although the pattern of uptake  $\beta$ -galactoside liposome is in essence similar in all the reported work by various investigators in relation to its preferential uptake by the parenchymal cells, there exists some difference as to the rate of uptake between our observations and that of Scherph of and Nicolou. In this context, it will be interest to mention that these investigators employed phosphatidyl serine containing lactosylceramide liposomes. It was reported by Hampton *et al.*<sup>30</sup> in their *in vitro* experiment that phosphatidyl serine containing liposomes does not bind with the galactose bringing lectin from castor bean unless Ca<sup>++</sup> present. In view of this *in vitro* effect, it may be

suggested that the phosphatidyl serine may have some effect on the receptor mediated uptake. This will explain the slow rate of uptake observed by these investigators. This point needs to be further investigated.

It is of interest to mention here that about 10 years ago we started with the concept that it should be possible to deliver drugs or enzymes intracellularly into lysosome in the case of diseases associated with lysosomal enzymes e.g. sphingolipidosis. It is apparent that we have been able to develop a very interesting model system using  $\beta$ -galactosylated liposomes to do this job (Fig. 2).

It may, however, be noted that it is yet to be ascertained about the stability of the external enzyme and expression of its activity after the delivery into the lysosomes. It is gratifying for me that it has been possible to extend our *in vitro* kinetic studies on lectin-liposome interaction to the use of glycolipid liposome or the targeting of specific liver cell types *in vivo* simply by varying the sugar residue on the surface of liposome.

#### *Oral Administration of Liposome*

In an earlier section, I have discussed *in vivo* administration of liposomes through intravenous route. In recent years attempts have been made to administer liposome entrapped material orally.<sup>29,30</sup> This route of administration has obvious advantages.

In a number of experiments, it was observed that orally administered liposome entrapped protein such as labelled gamma globulin was completely degraded when the radio activity was detected in cardiac blood.

Thus, these experiments were of inconclusive nature as it was not definite whether this degradation was in the intestine along with the disruption of liposome or in the organs. In later experiments, it was observed that a significant amount of liposome entrapped material could be detected in the portal blood. It was further observed that in the portal blood 75 per cent of the entrapped protein was intact and 50 per cent of the total gamma globulin in the portal blood was still intact as the liposome entrapped material. At the same time all the gamma globulin in the cardiac blood was found to be in the degraded form. This observation may be explained by assuming that liposome present in portal blood was taken up by the liver, processed there, and then circulated.

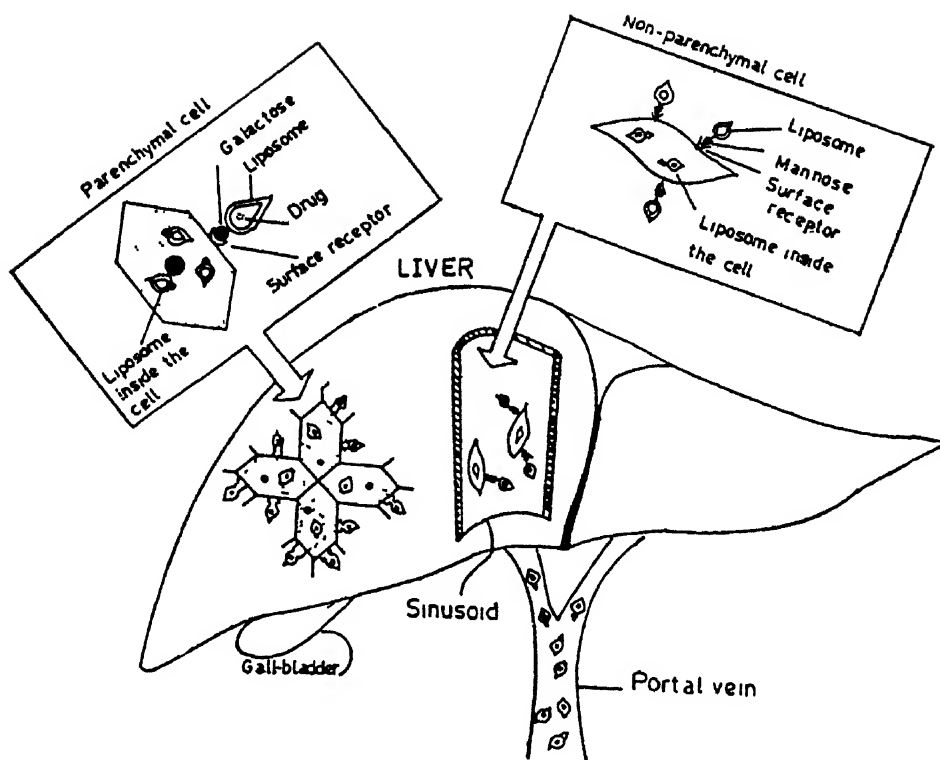


FIG 2 Targeting of drug to specific liver-cell types

When asialo GM<sub>1</sub> liposomes were employed there was a very rapid uptake by the liver. This uptake was actually inhibited by asialo fetuin administration at a suitable time interval. The presence of asialofetuin also increase the concentration of undergraded material in the portal blood. This was the first observation to show that orally administered liposomes enter into circulation through portal route. This study also indicates that although total amount of intact liposome is only 2-3 per cent of the administered dose, this mode of administration has a considerable potential as a drug delivery system. The surface modification of liposomes will also help in directing these liposomes even through oral route.<sup>33</sup>

It is hoped that in future it will be possible to increase the stability of liposome in the intestine. This will have an influence on the orally administered liposome entrapped drug.

Oral administration of liposome and its uptake by the liver suggests a rather exciting possibility with respect to indigenous Ayurvedic drug. It is well known that Ayurvedic drugs are effective when administered orally and quite a few of these drugs are in the form of glycosides. The terminal sugar of the drug and the way it is prepared and administered may have a

profound effect on the passage of this drug through intestine and uptake of the drug by various organs.

### C. LIPOSOME IN IMMUNOLOGY

The adjuvant effect of liposome was first reported by Allison and Gregoriaidis.<sup>34</sup> Since then the immunopotentiating effects of liposome have been the subject of intensive study. Liposome has an advantage over adjuvants currently used for human and animal immunization since they are prepared from biodegradable phospholipids which does to produce any granuloma. Further, liposomes protect the antigens from the hypersensitivity reaction.

In the laboratory of the Indian Institute of Chemical Biology, we have initiated a series of research work on the immunopotentiating effect of liposomes. Using protein as an antigen we have confirmed that liposome entrapped antigen also can be used when liposomes are employed as an adjuvant. We made the important observation that the nature of the surface of liposome has a significant effect on the level of antibody titre. It was observed that a lysozyme (a protein antigen) entrapped in neutral and negatively charged liposomes is higher than that with negatively charged or neutral liposomes or even with complete adjuvant. The strong immune response was found to be accompanied by mild granuloma<sup>37</sup> formation at the site of the injection. The exact mechanism of adjuvant action of the liposome is not known. There is a possibility that positively charged liposomes interact differently with cells *in vivo* in comparison to that of neutral and negatively charged liposomes. Similar immunopotentiating effect were also observed when synthetic phospholipids such as dipalmitoyl phosphatidyl choline and distearyl phosphatidyl-choline were used during the preparation of liposomes instead of egg lecithin. The best route of administration of liposome entrapped antigen to produce maximum antibody titre was found to be subcutaneous.

In order to understand the mechanism of action of liposome as an adjuvant, the vesicles were designed to specifically interact with macrophages, the phagocytic cell responsible for the clearance of liposomes. When antigen entrapped liposomes with mannose and galactose exposed on the surface were injected, liposome with terminal galactose residues induced an immune response comparable to adjuvant effect of sugar free neutral liposomes,<sup>38</sup> whereas the immune response of

mannosylated liposomes was equal to that of the free antigen. It is tempting to postulate that mannose-liposomes was equal to that of the free antigen. It is tempting to postulate that mannose-liposomes are taken by the macrophage *via* a receptor mediated process recognising the mannose residue resulting in the rapid degradation of the entrapped antigen and accordingly not available for eliciting antibody response.

#### D. LIPOSOME AS A HAPTENIC CARRIER FOR SMALL MOLECULES

In addition to their adjuvant effect, liposomes have been recognised as an efficient haptenic carrier of antigens in recent years. Haptenic groups are attached to liposomes by coupling them to phosphatidyl ethanolamine. With such liposomes dual properties of liposomes as carrier and adjuvant in eliciting antisaccharide (antigalactosyl and antimannosyl) immune response<sup>40,41,42</sup> were apparent.

Liposome mediated antigalactosyl antibody response when compared was found to be as good as either in the presence of complete Freund's adjuvant or through the conventional method of protein carrier.

This use of liposome as an haptenic carrier may be extended to various peptide fragments or small peptides. This will minimize the cross reactivity of the carrier protein.

The work presented in this review highlights the potential applications of liposomes in clinical science. This potential can be fully realized once we can develop stable unilamellar liposomes, stable *in vitro* and *in vivo*.

#### ACKNOWLEDGEMENTS

It is a great pleasure and pride for me to acknowledge the contribution of various colleagues and graduate students whose continuous inputs in this field had made our work a success. They are M K Das a Surolia, D Thambi Dorai, P Ghosh, P K Das, P Das Gupta, N Latif, N Das and D Sarkar. The Council of Scientific and Industrial Research and the Department of Science and Technology, Government of India provided the financial support which made this work possible. The author gratefully acknowledges the help of Dr Sandip K Basu during the preparation of this manuscript.

## References

1. A Surolia B K Bachhawat and S K Poddar *Nature* **257** (1975) 802
2. A Surolia and B K Bachhawat *Biochem Biophys Res Commun* **83** (1978) 779
3. A Syroliia B K Bachhawat and S K Poddar *Biomembrane Proc natn Symp Biol Memb Model System NIMHANS Bangalore* (Ed. S U Tablkov and S K Podder) (1978) 4
4. S K Podder A Surolia and B K Bachhawat *Indian J Biochem Biophys* **18** (1981) 322
5. C W M Grant and F W Peters *Biochem Biophys Acta* **779** (1984) 403
6. J H Austin A S Balasubramanian T N Pattabiraman S Saraswathi D K Basu and B K Bachhawat *J Neurochem* **10** (1963) 805
7. J H Austin D MAC Afee D Armstrong L Shearer and B K Bachhawat *Biochim J* **93** (1964) 15c
8. K A Balasubramanian and B K Bachhawat *Biochim Biophys Acta* **403** (1978) 113
9. G Gregoriadis P D Leathwood and B E Raman *FEBS Lett* **14** (1971) 95
10. G Ashwell and A G Morrell *Trends biochem Sci* **2** (1977) 76
11. P D Stahl J S Rodman M J Miller and P H Schlesinger *Proc natn Acad Sci (USA)* **75** (1978) 1399
12. B K Bachhawat *Neurology (India)* **22** (1974) 169
13. A Surolia and B K Bachhawat *Biochim Biophys Acta* **497** (1977) 760
14. P Ghose and B K Bachhawat *Biochim Biophys Acta* **632** (1980) 562
15. P Ghosh P K Das and B K Bachhawat *Arch Biochem Biophys* **213** (1982) 266
16. P Ghosh B K Bachhawat and A Surolia *Arch Biochem Biophys* **213** (1982) 266
17. G Gregoriadis and D E Neerunjan *Eur J Biochem* **47** (1974) 179
18. M M Jonah E A Ceruy and Y E Rahman *Biochem Biophys Acta* **541** (1978) 321
19. P Ghosh and B K Bachhawat *J Scient Ind Res* **39** (1980) 689
20. F C Szoka and E Mayhew *Biochem Biophys Res Comm* **110** (1983) 140
21. H H Spanger and C L Sherphof *Biochem Biophys Acta* **737** (1983) 40
22. D Keppler in *Pathogenesis and Mechanism of Liver Cell Necrosis* (Ed. Keppler) University Parle Press Baltimore MD (1975) 87
23. C Mathias D Thambi Dorai and B K Bachhawat *Indian J Biochem Biophys* **14** (1977) 142
24. P Ghosh P K Das and B K Bachhawat *Biochem Soc Trans* **9** (1981) 512
25. – *Proc natn Acad Sci India* **48A** (1982) 12



26. P Sorlan J Dijkstra A Legerand H Spanjer D Londs-Gagliandi F Roedrik, G Scherph of and G Nicolou *Proc natn Acad Sci (USA)* **80** (1983) 7128
27. P Dasgupta and B K Bachhawat *Biochem int* **10** (1985) 327
28. B K Bachhawat and P Das Gupta An approach to tissue targetting of drugs and proteins using glycolipid liposomes *Proc int Conf med Appl Liposomes Nagoya Japan* (1984) Dec 14-17
29. B K Bachhawat P K Das and P Ghosh In : *Liposome Technology, III* (Ed. G. Gregoriadis) CRC Press Inc Boca Raton FL (1983) 117
30. P Y Hapmton R W Holz and I J Goldstein *J biol Chem* **255** (1980) 6766
31. H M Patel and B E Raman *FEBS Lett* **62** (1976) 60
32. H C Henkeor W T H Hrmens A D Muller and R F A Zwaat *Lancet* **1** (1980) 70
33. N Das M K Das and B K Bachhawat *J appl Biochem* **6** (1984) 346
34. A C Allison and G Gregoriadis *Nature* **252** (1974) 252
35. P N Shek and B H Sabiston *Immunology* **45** (1912) 347
36. P N Shek B Y K Yong and N Z Stance *Immunology* **49** (1984) 122
37. N Latif and B K Bachhawat *Biosci Rep* **4** (1984) 99
38. – *Immunol Lett* **8** (1984) 75
39. —*J Biosci* **6** (1984) 491
40. P K Das P Ghosh B K Bachhawat and M K Das *Immunol Cooun* **11** (1983) 19
41. – *Experientia* **38** (1982) 629
42. D P Sarkar P K Das B K Bachhawat and M K Das *Immunol Commun* **11** (1982) 175



**Mahtab Sohrab Bamji** (b. 5 October 1934) did Ph.D.(1961) from University of Bombay. She is Scientist (Director's grade), National Institute of Nutrition (NIN), Hyderabad; In Charge, Laboratory Animals Information Service Centre, NIN.

Bamji developed a biochemical test for early detection of riboflavin deficiency and used it for assessing riboflavin-nutrition status of populations and riboflavin requirement. Elucidated the molecular basis of skin lesions, of riboflavin and pyridoxine deficiency (impaired collagen crossing)

Demonstrated the aetiological role of respiratory infections in riboflavin deficiency, and effects of these deficiencies on wound healing and psychomotor performance. Studied metabolic safety of hormonal contraceptives. Demonstrated the vitamin-like role of carnitine, which is synthesized from the essential amino acid lysine.

Bamji is Fellow of National Academy of Agricultural Sciences (India); Member, INSA Council (1993-95). She is the recipient of Shakuntala Devi Amirchand Prize (ICMR) (1971); Patwardhan Award (ICMR) (1983); Amrut Mody Award (1985); Dr. B.C Guha Memorial Lecture Award (INSA) (1987); G.J.S. Rao Memorial Lecture (IISc) (1992).

*Mahtab Sohrab Bamji was elected to the fellowship of the Academy in 1990.*

**Address:** Scientist (Director's grade), National Institute of Nutrition, Jamia Osmania P.O., Hyderabad 500 007; Tel: (0842) 868920; Fax: 91-842-869074; Grams: NUTRITION; Res: Apartment No. 211, Shree Datta Sai Apartments, 1-7-1074, RTC Cross Road, Musheerabad, Hyderabad 500 020; Tel: (0842) 615148.

## **BIOCHEMICAL DETECTION, AETIOLOGY AND FUNCTIONAL CONSEQUENCES OF RIBOFLAVIN DEFICIENCY**

MAHTAB S BAMJI

*The erythrocyte glutathione reductase test reveals a very high incidence (over 80%) of riboflavin deficiency among low-income group populations in and around Hyderabad 15-20% show clinical lesions of riboflavin deficiency, such as mucocutaneous lesions of the mouth Besides dietary deficiency, urinary losses during diseases such as upper respiratory infections and measles may also contribute to the high incidence of riboflavin deficiency Skin lesions of riboflavin and pyridoxane deficiency in animals and humans are similar and may have a common molecular basis viz impaired collagen cross-linking Riboflavin deficiency impairs pyridoxal phosphate (PLP) synthesis and PLP is essential for collagen cross-linking Hitherto unrecognized functional consequences of riboflavin deficiency are, delayed wound healing and reduced hand steadiness*

### **BIOCHEMICAL DETECTION OF RIBOFLAVIN DEFICIENCY**

Several years ago we developed a test for the assessment of riboflavin nutrition status, based on the *in vitro* activation of the flavin enzyme, erythrocyte glutathione reductase by its coenzyme FAD (Bamji 1969). The logic in this test is, in riboflavin deficiency, the enzyme activity drops due to reduction in the levels of the coenzyme FAD. However, the enzyme activity can be restored by *in vitro* addition of FAD. *In vitro* activation of glutathione reductase referred to as EGR-AC is inversely related to riboflavin deficiency. Following guidelines for interpreting this test were derived from control depletion-repletion studies, and values seen in subjects suffering from riboflavin deficiency. EGR-AC < 1.2, normal, 1.2-1.4, marginal deficiency > 1.4, deficient. Subsequent surveys of population groups, particularly women and school children show a very high incidence of biochemical riboflavin deficiency in and around Hyderabad (Bamji et al. 1979, Bamji et al. 1982, WHO task force on oral contraceptives 1986, Prasad et al. 1987) (table I).

## AETIOLOGY OF RIBOFLAVIN DEFICIENCY

Diet surveys conducted by the National Nutrition Monitoring Bureau (ICMR) show that vitamin A and riboflavin are among the most limiting nutrients in Indian diets (National Nutrition Monitoring Bureau Report 1981). Dietary riboflavin deficiency is particularly marked in the southern and eastern states of India where rice is the staple diet. Situation improves with a mixed cereal millet diet (table 2). Poverty, ignorance, wrong food preferences and cooking practices all contribute to dietary riboflavin deficiency. We feel that apart from diet, infections also play a significant role in development of riboflavin deficiency. Very often normalization of EGR-AC does not occur despite administration of riboflavin supplements. A study in rural children showed an association between seasonal rise in upper respiratory infections and urinary excretion of riboflavin during winter (Sarma et al. 1981). Subsequent study in urban preschool children showed that during measles and bouts of upper respiratory infections (URI) there is a marked rise in the urinary excretion of riboflavin (Bamji et al. 1987) (table 3). Such metabolic losses may perpetuate riboflavin deficiency and may account for its very high incidence. Similar phenomenon does not seem to operate for other B- complex vitamins. Along with rise in urinary excretion, a transient rise in blood levels also occur suggesting mobilisation of riboflavin from the tissues.

Table 1

*Prevalence of riboflavin deficiency as judged by erythrocyte glutathione reductase activation coefficient (EGR-AC) among low-income group women and children*

| EGR-AC                               |        |                                   |                          |                        |         |  |
|--------------------------------------|--------|-----------------------------------|--------------------------|------------------------|---------|--|
|                                      | Number | Percentage frequency distribution |                          |                        |         | Source   |
|                                      |        | 1.20<br>Low risk                  | 1.21-1.40<br>Medium risk | 1.41-1.60<br>High risk | > 1.610 |  |
| Rural children (1-5 years)           | 105    | 21                                | 33                       | 21                     | 25      | Unpublished  |
| Rural school boys,<br>5-11 years     | 114    | 0                                 | 52                       | 155                    | 793     | Bamji et al. 1982  |
| Urban school children,<br>7-11 years | 103    | 2                                 | 4                        | 9                      | 85      | Prasad et al. 1987   |
| Rural women<br>15-45 years           | 105    | 8                                 | 21                       | 21                     | 50      | Unpublished  |
| Urban Women<br>18-35 years           | 415    | 8                                 | 10                       | 14                     | 68      | Bamji, Prema &<br>Jacob (WHO task<br>force,<br>unpublished |

**Table 2**  
*Riboflavin intake in Rural India\* % RDA for 2400 calories/C unit+*

|                | Riboflavin<br>intake<br>% RDA | Rice as % of<br>total calories |
|----------------|-------------------------------|--------------------------------|
| Kerala         | 64                            | 97                             |
| Tamil Nadu     | 63                            | 98                             |
| Orissa         | 52                            | 97                             |
| Andhra Pradesh | 59                            | 88                             |
| West Bengal    | 65                            | 86                             |
| Karnataka      | 73                            | 65                             |
| Gujarat        | 92                            | 38                             |
| Uttar Pradesh  | 101                           | 27                             |

\* Data from National Nutrition Monitoring Bureau (1981)

+C unit = Consumption unit For an adult man during sedentary work, C unit is 1. For others a scale of coefficients worked out based on calorie requirements.

**Table 3**  
*Effect of infections on urinary riboflavin levels in children*

|                              | Sample<br>size | Riboflavin mg/g<br>creatinine |         |
|------------------------------|----------------|-------------------------------|---------|
| Control                      | 21             | 0.39                          | 0.17    |
| Upper respiratory infections | 29             | 1.20                          | 0.73*** |
| Measles                      | 28             | 1.24                          | 0.15*** |

\*\*\* 0.001 compared to controls by 't' test

Certain drugs are also known to alter the riboflavin requirement. Our studies show that high dose formulations of oral contraceptives increase the riboflavin requirement due to selective increments in specific flavin enzymes. This leads to an unequal distribution of riboflavin between its enzyme systems creating pockets of deficiency and excess—a state of relative deficiency (Ahmed et al. 1975, Ahmed & Bamji 1976a, b).

## FUNCTIONAL CONSEQUENCES OF RIBOFLAVIN DEFICIENCY

Figure 1 describes the stages in the development of vitamin deficiency. The molecular basis of most vitamin deficiency diseases is not understood. There are reasons to believe that even at the subclinical stage of malnutrition some subtle functional impairments like changes in psychomotor performance may occur. Even an apparently healthy individual may not be in a state of optimum health.

Subclinical malnutrition can be detected only through appropriate biochemical tests. Our experience shows that while majority of the subjects who show clinical signs and symptoms of vitamin deficiency, show evidence of biochemical deficiency, within the biochemically deficient individuals, there are many who do not show clinical evidence. This is because, disease is a complex phenomenon in which factors other than vitamin deficiency may play a role in tipping an individual from subclinical to the clinical state (figure 1).

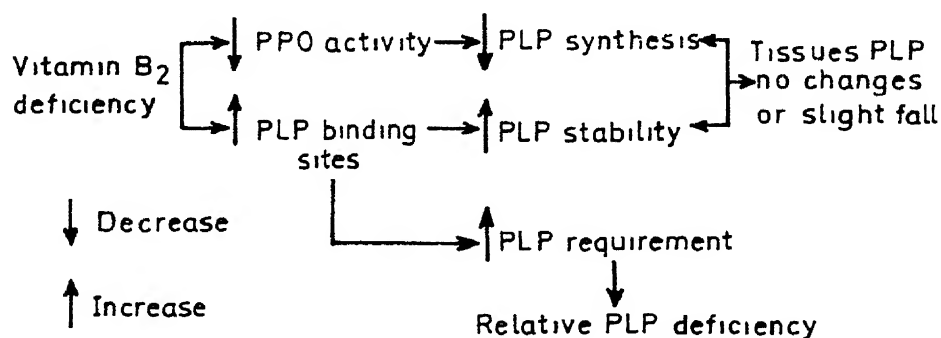
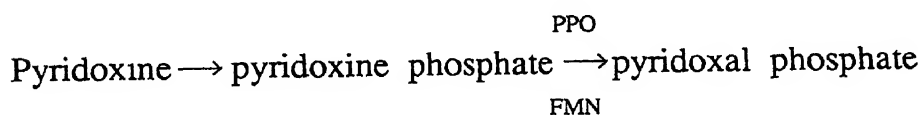


FIG 1 Development of nutritional deficiency disease

Flavin enzymes are essential for biological oxidation reduction and energy transduction reactions. Riboflavin deficiency in animals leads to loss of appetite, skin changes, growth failure and death. However, in human situation, where only sub-total deficiency occurs, the clinical morbidity, or riboflavin deficiency is not as serious as that of some of the other vitamins. The presenting clinical symptoms of riboflavin deficiency in humans are, mucocutaneous lesions of the mouth such as angular stomatitis, glossitis and cheilosis. One possible explanation for the less serious consequences of sub-total riboflavin deficiency as compared to the deficiencies of other B-complex vitamins in humans may be the fact that

important mainstream flavoproteins concerned with energy metabolism do not lose their activity till very terminal stage of deficiency because the dissociation constants ( $K_d$ ) for their coenzymes are very low (Burch et al. 1956). It is therefore possible that the pathology of riboflavin deficiency is due to molecular event(s) related to side-stream reactions and hence are of less serious consequences.

*Molecular basis of the skin lesions of riboflavin deficiency.* Several years ago some of my clinical colleagues observed that the oral lesions regarded to be pathognomic of riboflavin deficiency, sometimes do not respond fully to treatment with riboflavin. However, they respond to treatment with pyridoxine (Krishnaswamy 1971, Iyengar 1973). Experimental deficiencies of vitamins B<sub>2</sub> as well as B<sub>6</sub> have also been reported to produce skin changes in rats and mucocutaneous lesions of the mouth such as angular stomatitis, glossitis, cheilosis and masolabial dyssebasia in human (Goldsmith 1962, Viter 1962). This raised the possibility of unified hypothesis explaining the roles of these two vitamins in the development of skin lesions. Wada and Snell (1961) have reported that the conversion of pyridoxine to its coenzyme pyridoxal phosphate depends on the FMN-dependent enzyme, pyridoxaminephosphate oxidase (PPO) in the following reaction sequence:



Based on this, we postulated that the biochemical basis of the skin lesions which respond to riboflavin and/or pyridoxine maybe cellular deficiency of pyridoxal phosphate (PLP) caused either by dietary deficiency of pyridoxine *per se* or secondary to impaired synthesis of PLP due to riboflavin deficiency. Since riboflavin is the more limiting vitamin in the diet than pyridoxine, these lesions are normally observed to respond to treatment with riboflavin.

To substantiate this hypothesis we have examined the effects of riboflavin deficiency in rat and humans on the metabolism of pyridoxine to pyridoxal phosphate. These studies showed that in riboflavin deficiency the activity of flavin enzyme pyridoxaminephosphate oxidase is markedly reduced and synthesis of PLP impaired (Lakshmi & Bamji 1974, 1976, 1979) (tables 4,5,6). Additionally it was observed that riboflavin deficiency also raises the cellular requirement of pyridoxine by increasing

the levels of some PLP-dependent enzymes (Lakshmi & Bamji 1975). This leads to a bizarre redistribution of PLP between its enzyme systems creating pockets of deficiency and excess i.e. state of relative deficiency (figure 2).

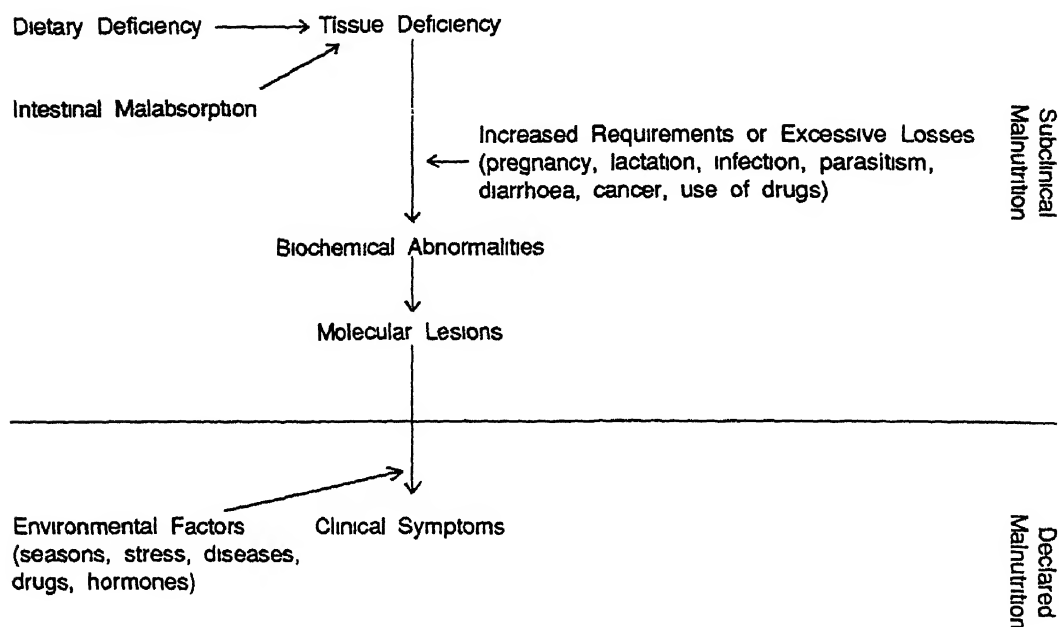


FIG 2 Interrelationship between riboflavin and pyridoxine (vitamins B<sub>2</sub> and B<sub>6</sub>)  
Source : M S Bamji, 1986

**Table 4**  
*Effect of riboflavin deficiency on pyridoxal phosphate (PLP)  
synthesis in rats*

|   | Control    | Deficient  |
|---|------------|------------|
| Liver PPO Activity*   | 165 ± 12.1 | 27.5 ± 4.3 |
| <i>In vivo</i> incorporation of<br><sup>14</sup> C pyridoxine into liver PLP<br>(% total radioactivity) |            |            |
| 3 min after injection   | 28.20      | 6.07       |
| 10 min after injection  | 37.13      | 1.35       |

Pyridoxaminephosphate oxidase activity expressed as PLP formed from pyridoxine phosphate  
(n mol)/g liver per 30 min  
Mean ± SE



**Table 5**  
*Effect of riboflavin deficiency on in vivo PLP synthesis in humans*

| No. of subjects | Clinical signs | Biochemical deficiency | PLP ratio |
|-----------------|----------------|------------------------|-----------|
| 4               | absent         | absent                 | 8.25±0.72 |
| 2               | absent         | present                | 3.94±5.03 |
| 10 (BT)         | present        | present                | 3.61±0.76 |
| 7 (AT)          | absent         | absent                 | 5.60±1.02 |

PLP ratio= Blood PLP concentration, 15 min after parenteral administration of 5 mg pyridoxine/blood PLP before pyridoxine administration  
 BT before treatment; AT, after treatment with 5 mg riboflavin for 10 days

**Table 6**  
*Effect of riboflavin deficiency on in vitro synthesis of PLP in human erythrocytes*

| Substrate     | Pyridoxine<br>µg PLP formed/ml | Pyridoxaminephosphate<br>RBC in 30 min. |
|---------------|--------------------------------|---|
| Controls (6)  | 0.76±0.03                      | 0.69±0.11                               |
| Deficient (7) | 0.19*±0.03                     | 0.20*±0.02                              |

\*P 0.05 compared to controls

( ) Number of subjects

The next question was, why should absolute or relative PLP deficiency cause skin lesions. PLP was suggested to be a cofactor for the enzyme lysyl oxidase which initiates the process of collagen cross linking (Murray & levine 1971). Collagen is an important connective tissue protein present in the dermis of the skin. Based on this we hypothesized that the molecular basis of skin lesions in riboflavin deficiency is impaired collagen crosslinking. This would weaken the dermis of the skin, and render the overlying epithelial tissue susceptible to stress and infections. The lesions would be more marked in the areas of constant stress and friction such as the mucocutaneous junctions. Normal collagen is also essential for epidermal growth and differentiation (Bano et al. 1983). Our experiments in rats show that riboflavin as well as pyridoxine deficiency impair cross-linking in skin collagen as judged by several biochemical,

1986a) (tables 7,8). Thus both vitamins B<sub>2</sub> and B<sub>6</sub> deficiencies reduce skin collagen content and aldehyde levels, increase collagen solubility in non-polar solvents such as NaCl and in denaturing agents, such as urea and potassium thiocyanate. They increase collagen gel reversibility and decrease its tensile strength and shrinkage temperature. All these changes suggest that deficiencies of these vitamins affect collagen cross-linking. Studies with <sup>3</sup>H-proline suggest that collagen synthesis is also affected (Prasad et al. 1986b).

**Table 7**  
*Skin collagen content and properties in vitamins B<sub>2</sub> and B<sub>6</sub> - deficient growing male rats*

|  | Food restricted<br>(weight matched control) | Vitamin B <sub>2</sub> deficient | Vitamin B <sub>6</sub> deficient |
|--|---|----------------------------------|----------------------------------|
| Percentage of <i>ad lib</i> - control rats |   |                                  |                                  |
| Total collagen                             | 83.1  | 72.1                             | 78.3                             |
| Insoluble collagen                         | 80.7  | 67.4                             | 70.3                             |
| Aldehyde content                           | 94.9  | 70.8                             | 61.8                             |
| $\alpha/\beta$ subunit ratio               | 130   | 142                              | 143                              |
| Tensile strength                           | 33.2  | 22.2                             | 26.5                             |

**Table 8**  
*Effect of riboflavin or pyridoxine deficiency on susceptibility of insoluble collagen to denaturing and proteolytic agents*

| Solubility of insoluble collagen by treatment with | Weight control | Riboflavin deficient | Pyridoxine deficient |
|--|----------------|----------------------|----------------------|
|  | (% of total)   |                      |                      |
| 6M urea  | 44.65          | 71.32                | 72.93                |
| 2M KCNS  | 41.83          | 76.46                | 74.46                |
| 0.0016% pronase                                    | 39.48          | 70.92                | 66.45                |

The above-mentioned experimental observations support our hypothesis that the molecular basis of skin lesions in riboflavin and

pyridoxine deficiency is impaired dermal collagen structure. Food restriction has similar (though not identical) effects but of lesser magnitude.

### ADDENDUM

Recent observations of Williamson and Kegan (1987) question the claim of earlier investigators that PLP is a cofactor for lysyl oxidase. Thus the mechanism of altered collagen properties in the deficiency of vitamins B<sub>2</sub> or B<sub>6</sub> remains to be elucidated. According to one possibility, increased levels of sulphur-containing amino acids like homocysteine in vitamin B<sub>6</sub> deficiency may interfere with collagen cross-link synthesis by binding to the lysine-derived aldehyde groups.

*Wound healing.* Collagen synthesis is an important event in the process of wound healing. Our recent unpublished observations show that healing of both excised and incised wounds is affected in the deficiency of riboflavin and pyridoxine (Laxmi A V, Laxmi R and Bamji M S, unpublished). Whether impaired collagen synthesis is the only factor responsible or other molecular events are also involved remains to be examined.

*Psychomotor function:* In two separate studies among rural (Bamji et al. 1982) and urban (Prasad, Lakshmi and Bamji unpublished) children we have observed that riboflavin deficiency impairs hand steadiness. Other psychomotor, performance tests such as cognitive ability, reaction time, finger dexterity (Bamji et. al 1982) and physical work capacity as judged by time taken to run a fixed distance and pulse rate after work, (unpublished observations) are not affected. Hand steadiness is an extrapyramidal function, dependent on dopamine activity. PLP is an essential cofactor for dopamine synthesis. In view of the earlier mentioned relationship between riboflavin deficiency and PLP synthesis it is possible that dopamine synthesis is impaired in riboflavin deficiency and this leads to subtle extrapyramidal deficits.

### References

1. Ahmed F, Bamji M S and Iyengar L 1975 Effect of oral contraceptive agents on vitamin nutrition status; *Am J. Clin. Nutr* 28 606  
— and — 1976a Biochemical basis of riboflavin defect associated with the use of oral contraceptives; *Contraception* 14 297

- and — 1976b Vitamin supplements to women using oral contraceptives; *Contraception* 14 309
2. Bamji M S 1969 Glutathione reductase activity in red blood cells and riboflavin nutrition status in humans; *Clin. Chem. Acta* 26 263  
 —, 1986 in *Nutrition and the Skin* pp 99-115 ed D A Roe (New York: Alan R Liss, Inc.)  
 —, Arya S, Sarma K V R and Radhaiah G 1982 Impact of Longterm low-dose vitamin supplements on vitamin status and psychomotor performance of rural school boys; *Nutr. Res.* 2 147  
 —, Bhaskaram P and Jacob C M 1987 Urinary riboflavin excretion and erythrocyte glutathione reductase activity in preschool children suffering from upper respiratory infections and measles; *Annl Nutr. Metab* 31 191  
 —, Rameshwar Sarma K V and Radhaiah G 1979 Relationship between biochemical and clinical indices of B vitamin deficiency – A study in rural school boys; *Brit J Nutr* 41 431
  3. Bano M, Zwiebel J A, Salmon D S and Kidwell W R 1983 Detection and Partial characterization of collagen synthesis stimulating activities in rat mammary adenocarcinomas; *J Biol Chem.* 258 2729
  4. Burch H B, Lowry O H, Padilla A M and Combs A M 1956 Effects of riboflavin deficiency and realimentation on flavin enzymes of tissues; *J Biol Chem* 223 29
  5. Goldsmith G 1962 Riboflavin malnutrition, in *Clinical Nutrition* ed N Jollite (New York: Hoeber Medical Diver Harper and Row) pp 604-621
  6. Iyengar L 1973 Oral lesions in pregnancy; *Lancet* 1 680
  7. Krishnaswamy K 1971 Erythrocyte transaminase activity in human vitamin B<sub>6</sub> deficiency; *Int J Vit Nutr. Res* 41 240
  8. Lakshmi A V and Bamji M S 1974 Tissue pyridoxal phosphate concentration and pyridoxamine phosphate oxidase activity in riboflavin deficiency in rats and man; *Brit. J Nutr* 32 249  
 — and — 1975 Pyridoxal-5' phosphate-dependent enzymes in riboflavin deficiency; *Ind. J. Biochem. Biophys* 12 136  
 — and — 1976 Regulation of blood pyridoxal phosphate in riboflavin deficiency in man; *Nutr. Metab* 20 228  
 — and — 1979 Metabolism of 2-<sup>14</sup>C pyridoxine in riboflavin deficiency; *Biochem. Med* 22 274
  9. Murray J C and Levene C I 1971 Evidence for the role of Vitamin B<sub>6</sub> as a cofactor for lysyl oxidase; *Biochem J.* 167 463
  10. National Nutrition Monitoring Bureau 1981 Annual Report for the year 1980 National Institute of Nutrition (ICMR), Hyderabad, India

11. Prasad P A, Lakshmi A V and Bamji M S 1987 Riboflavin and haemoglobin status of urban school boys; relationship with income, diet and anthropometry; *Ind. J. Pediat.* **54** 529
12. Prasad R, — and — 1983 Impaired collagen maturity in vitamin B<sub>2</sub> and B<sub>6</sub> deficiency – probable molecular basis of skin lesion; *Biochem. Med.* **30** 333  
 —, — and Venkatappaiah V 1986a Biochemical and biophysical properties of skin collagen in riboflavin and pyridoxine-deficient rats; *Clin. Biochem. Nutr* **1** 39
13. Sarma K V R, Radhaiah G and Bamji M S 1981 Impact of long-term, low-dose B-complex vitamin supplements on clinical and anthropometric status of rural school boys; *Nutr. Reports Internat.* **24** 345
14. Viter R W 1962 Vitamin B<sub>6</sub> (pyridoxine); in *Clinical Nutrition* Ed N Jolliffe 691 (New York: Howber Medical Division, Harper and Row)
15. Wada H and Snell E E 1961 The enzymatic oxidation of pyridoxine and pyridoxamine phosphate; *J Biol. Chem* **236** 2089
16. WHO Task Force on Oral Contraceptive 1986 Impact of hormonal contraceptives vis-a-vis non-hormonal factors on the vitamin status of malnourished women in India and Thailand; *Human Nutritn Clin. Nutr* **C40** 205
17. Williamson P R and Kagan H M (1987)  $\alpha$ -proton abstraction and carbanion formation in the mechanism of action of lysyl oxidase; *J. Biol Chem* **262** 8196



**Nihal Kishinchand Notani** (b. 21 January 1930) did Ph.D. (1957) from Purdue University, USA. He is INSA Senior Scientist (1990-93); Director, Biomedical Group, Bhabha Atomic Research Centre, Bombay; and was formerly FAO/IAEA Fellow, University of Cornell and University of Wisconsin; National Science Foundation (USA); Senior Foreign Scientist Fellow; Visiting Professor, University of Tennessee, USA.

Notani has shown the mechanism of transformation in bacterium *Haemophilus influenzae* to be through integration of single-stranded segments of input DNA into the homologous regions of resident chromosome; he has cloned two DNA repair genes; isolated and characterized a large, linear plasmids from *Streptomyces rhimosus*; analyzed a genetically unstable region in *Streptomyces lividans*; proposed an excision mechanism for maize transposable elements; transferred 3 DNA sequences including bar gene (resistance to herbicide Bialaphos) to tobacco, obtained true-breeding transgenic (herbicide-resistant) plants.

Notani is Fellow of Indian Academy of Sciences, National Academy of Sciences (India), Maharashtra Academy of Sciences, and National Academy of Agricultural Sciences (India); Elected Member, Human Genome Organization; has been Member of INSA Council (1987-89). He is the recipient of B.C. Guha Memorial Lecture Award (INSA) (1990).

*Nihal Kishinchand Notani was elected to the fellowship of the Academy in 1976.*

**Address:** Emeritus Scientist, Biomedical Group, Bhabha Atomic Research Centre, Trombay, Bombay 400 085; Tel: (022) 5510323; Telex: 011-71017 BARC IN; Fax: 91-22-5560750; Res: 43, Juhu Ajanta Cooperative Housing Society, Gulmohar Road, JVPD Scheme, Bombay 400 049; Tel: (022) 6204185.

## GENETIC TRANSFORMATION

N K NOTANI

I am thankful to the President, Indian National Science Academy for asking me to deliver the Bires Chandra Guha Memorial Lecture. This award is given for a distinct contribution in the fields of Biochemistry, etc. in the broadest sense. I chose the subject of Genetic Transformation because I have been working on it for more than two decades. I should justify my working for so long a period in this field by quoting Alfred Hershey who when asked about his idea of happiness said something like this : Happiness is an experiment that works and you keep doing it again and again. So, indeed my case is also that of familiarity breeds content. Moreover, transformation is still an active field and in fact now has been extended from bacteria to higher organisms.

What I thought I will do is to tell you not only the story of genetic transformation but also wherever possible to emphasize the difficulties in deciphering the mechanisms of sub-microphenomena which cannot be perceived by our senses directly, implying that logic-based models may not necessarily be true representation of the reality. Secondly, I would also like to talk about a certain dead-end kind of feeling that one gets about science at times, that it may be finished and there may be *ne plus ultra* (no more beyond). This question is not so antiheroic as may seem to evangelists of science, and already in 1969 Gunther Stent had written a book which was a view of the end of progress. Although, Stent's 'obituary' of science turned out to be premature, Peter Medawar explored the question in a different way in his 1984 book, "The Limits of Science". Medawar discussed his arguments in the context of Francis Bacon's 17th Century treatise, that there may be "no-more-beyond" because of intrinsic limits on indefinite growth of anything (say, skyscraper or a bacterial colony) or loss of nerve or cognitive inadequacy. Bacon as you know was also the prime inspiration in the start-up of the Royal Society of London as a contemporary engraving makes clear with the Society President sitting on the right of the bust of King Charles II and Bacon on his left (see Bernal 1969). Sir Francis is also credited with what has come to be called the Baconian creed which exhorts scientists to make applications from their work. The long-range ramifications of genetic transformation have led to the production of transgenic plants and animals on the one hand and

experiments on gene therapy in man on the other. Therefore, I will also briefly cover the work on transgenic plants exemplifying research done at BARC and other Laboratories.

## GENETIC TRANSFORMATION IN BACTERIA WITH HOMOLOGOUS DNA

Genetic transformation was discovered by F Griffith in 1928. He noted that when either heat-killed pathogenic pneumococci or live non-pathogenic pneumococci were injected into mice, the animals survived. However, if a mixture of the two was injected it caused bacteraemia and the mice died. Apparently, something from the dead (pathogenic) bacteria was getting transferred to live bacteria turning them into pathogens. Later, these experiments could be done *in vitro*. A cell-free extract from pathogenic bacteria could effect a heritable change in the non-pathogenic bacteria. The cell-free extract containing the so-called Transforming Principle (TP) was analyzed by Avery et al. 1944. Their findings were unexpected; TP was not a protein but DNA (Deoxyribonucleic acid). This observation was somewhat ahead of its times and because it could not be connected to the canonical knowledge (Stent 1978), its general validity and acceptance took some more time. As Hershey (1966) has put it that some redundancy of evidence was needed and some diversity of experimental material was crucial.

With the elucidation of double-helical structure of DNA by Watson and Crick, plausible molecular hypotheses for major genetical functions could be formulated. For example, two major alternate possibilities were considered for mechanism of genetic recombination: one was the so-called Breakage-rejoin and the other Copying-choice. Apart from pneumococci, two other bacterial species viz. *Haemophilus influenzae* and *Bacillus subtilis* were also shown to be naturally transformable. We tested these hypotheses in *H. influenzae* system and showed that the transformation is effected through a kind of breakage-join mechanism i.e. a single-stranded segment of donor DNA displaces its homologous counterpart in the resident DNA (Notani & Goodgal 1966). Since these experiments were done with the two DNAs (donor and resident) differentially marked for genetic alleles and density labels, the primary recombinant molecule in the transformed region was inferred to be hybrid for the density label and heterozygous for the genetic marker. Similar results were obtained with pneumococci (Fox & Allen 1964, Fox 1966).



In pneumococci, a single-stranded donor DNA intermediate had been discovered (Lacks 1962). However, no such free single-stranded form of donor DNA was detectable in *H. influenzae*. Instead, slower-sedimenting fragments with poor transforming activity were detected but these appeared to be *byproducts* of transformation (Notani 1971).

Jane Setlow had isolated two recombination-deficient mutations designated *rec1* and *rec2*. Although, uptake of donor DNA was normal in the two mutant strains, the transformation was down by five to six orders of magnitude. Our genetic analysis indicated that the two mutations are in different loci.

Molecular analysis showed that *rec2* gene controlled an early step and *rec1* a later one. In *Rec2*<sup>-</sup> strain, homologous DNA was taken up normally but failed to integrate into the resident DNA. Donor DNA remained intact for a very long time and showed no signs of any degradation. On the other hand, in *Rec1*<sup>-</sup> strain, there was considerable intracellular degradation of donor DNA. These products of degradation were incorporated into the resident DNA. No genetic information was transferred (Notani et al. 1972).

At this stage the story of genetic transformation in haemophilus seemed complete in its outline. Was there anything more beyond (*plus ultra*) ? For a while it seemed there was nothing exciting but then Hamilton Smith showed that haemophilus in the state of competence takes up its own DNA *preferentially* and that the basis for this is the presence of an 11-bp recognition and uptake sequence, 600 copies of which are distributed in its genome. Thus, every few kb there would be an uptake sequence and any fragment then would have a good chance of getting taken up by *H. influenzae* (competent) cell.

Another new observation was that in the state of competence, haemophilus sprouts membranous outgrowths termed transformasomes (Kahn et al, 1982, 1983). Immediately following uptake, donor DNA is localized and sequestered there, where it is protected from the action of external DNase and cellular restriction enzymes. DNA then exits through a 'pore' into the cell and recombines with the resident DNA (Barany et al. 1983).

## GENETIC TRANSFORMATION WITH PLASMID AND CHIMERIC PLASMID DNA

In 1979, we started to work with a plasmid RSF0885 which carried an ampicillin resistance marker ( $amp^r$ ) and which plasmid had been isolated from a natural strain of haemophilus. Transformation with this plasmid was very inefficient (Notani et al. 1981). This was a puzzling feature but then we interpreted this observation in terms of RSF0885 lacking uptake sequences. Predictably, when we spliced chromosomal segments to it, the transformation frequency went up by two to three orders of magnitude (Setlow et al. 1981, Notani 1981). It also provided a facile method for detecting inserts.

Another puzzling observation was that transformation with chimeric plasmid DNA which is fixed *extrachromosomally*, still needed the recipient's *rec* gene expression for efficient transformation. We proposed a model which implicated recombination to translocate incoming chimaeric DNA into the cell (Joshi et al. 1984, figure 1).

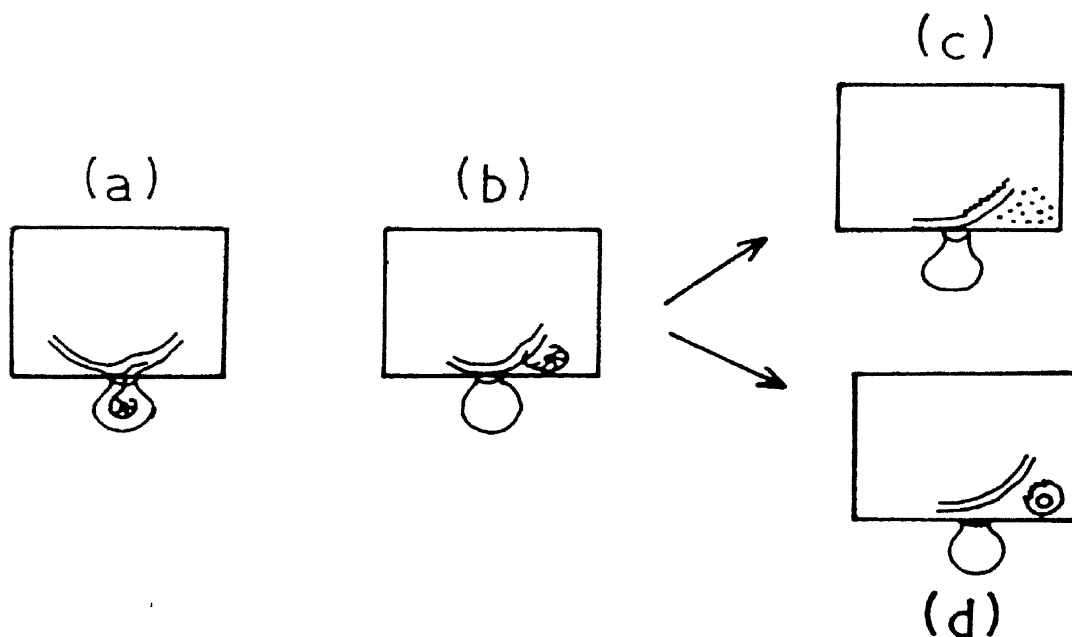


FIG 1 Imagined scheme for uptake and extrachromosomal fixation of chimaeric plasmid DNA taking into account the role of recombination in *H. influenzae*  $nov^r$  marks the chromosomal insert and  $amp^r$  marks the vector plasmid. (a) Chimeric plasmid enters transformasome and one-strand then exits through the 'pore' It pairs with the homologous segment in the chromosome at the nexus; (b) As the chromosome rotates away, the plasmid is brought in, (c) Only the insert is integrated in the resident DNA and rest of the plasmid breaks down or, (d) Plasmid is fixed extrachromosomally yielding  $amp^r$  transformant. Chromosome marker, if exchanged, would be  $nov^s$ .  $Amp^r$  transformants are no more than a few percent, indicating most of the time plasmid is degraded

This imagined hypothesis was tested using a mutant strain N19 which had been isolated in our laboratory a long time ago. A special feature of N19 strain is that it yields *nov* transformants with one hundredth the frequency of that obtained from the wild type strain(17). Apparently, it has an aberration in *nov* region. We found that extrachromosomal fixation of *nov* gene-carrying plasmids is much less efficient in N19 than say of a clone carrying *str<sup>r</sup>* marker providing support for our idea that recombination *per se* is involved in translocating and fixing chimaeric DNA, extrachromosomally (Joshi & Notani 1988).

Self-cloning of genes was working very efficiently specially with new vectors that were constructed in our laboratory by Dr (Mrs) V P Joshi. Not only several antibiotic-resistance alleles of genes like *nov* and *str* been cloned but also two DNA repair genes *uvr* (Kanade & Notani 1987) and *uvr<sup>3</sup>* (Mody et al. 1990). Also, one can clone very large segments of DNA (upto 36kb) with a plasmid vector in *H. influenzae* system (V P Joshi, oral communication).

## GENETIC TRANSFORMATION OF PLANTS

Although, genetic transformation of plants was attempted initially using bulk DNA, the results were not completely convincing. This had to await the development of *Agrobacterium tumefaciens* gene-transfer system. *A. tumefaciens* infecting dicotyledonous plants through wounds induce tumours. The basis of tumorigenesis is the transfer of a segment of DNA (called T-DNA) from its plasmid to plant genome following infection. If any DNA segment is inserted between the two ends of T-DNA, it will also be transferred to plant genome. The development of this system is primarily due to the efforts of Jeff Schell and Marc van Montagu. Viegas in our lab has adopted this technology which requires the following steps:

(i) Selection and isolation of a gene, usually from a prokaryote or cDNA of a eukaryote and making a DNA construction with appropriate expression signals that will express the gene to be transferred in plants, (ii) Cloning of this construction in *E. coli* and its mobilization from it into agrobacteria, (iii) Infection of plant leaf discs or cotyledons or protoplasts with agrobacteria carrying the DNA construct. (iv) Selection of transformed leaf discs or protoplasts and growing them in tissue culture and finally to maturity, (v) Segregation of the transferred marker and its detection in the next generation.

Elsewhere several gene transfers have been made which has enabled the production of insect, herbicide and viral disease-resistant plants. At BARC two gene transfers have been made to tobacco which characters have been transmitted vertically. Initially we made only a gene marker transfer and we learnt that more than one copy of T-DNA can be transferred (Viegas et al. 1987). Then we made *nif* HD DNA transfer without putting the plant expression signals and found that even without selection such transfers can be made. Also, *nif*HD DNA, remained linked to *kan<sup>r</sup>* marker in about 70% of the cases (Viegas & Notani 1988).

A crystal protein gene has been isolated from *Bacillus thuringiensis* and cloned in *E. coli*. This gene is expressed very well in *E. coli* and the crystal is visible in the transgenic bacteria under the light microscope. An extract of it has been shown to have lepidopteran larvicidal activity (Tuli et al. 1989).

A fourth gene, *bar*, which imparts resistance to herbicide bialaphos has been subcloned in *E. coli* and mobilized into agrobacterium. The bacteria were used to infect tobacco leaf discs and some transformants resistant to herbicide bialaphos have been obtained (Viegas, unpublished).

While transformation of dicotyledonous plants has now become almost routine, transformation of monocot cereals is not easily obtained. In fate experiments, we found that input DNA imbibed with wheat and rice seeds or germs, is broken down quite rapidly, providing evidence for the presence of external and internal nucleases (Viegas & Notani 1990).

## DISCUSSION

Main features of genetic transformation of bacteria appear to be reasonably worked out but there seems to be 'more beyond' (*plus ultra*). For example, the interaction between input chimaeric plasmid DNA and resident DNA is quite remarkable but so far its analysis has been purely genetical. DNA analysis may throw further light on various steps in fixation of chimaeric DNA. Hot spots of recombination are another characteristic of certain segments of DNA which needs further analysis. A property of haemophilus system is that, very large segments of DNA with a plasmid vector can be cloned, although expression in this system is not worked out. Cloning of any DNA specially eukaryotic remains an attractive possibility in this system.

As far as transformation of dicotyledonous plants is concerned, sky seems to be the limit. All kinds of genes are being transferred in all kinds

of crops. Starting with single gene transfers for resistance to herbicides, insect pests and viruses; large-scale transfers of sequences corresponding to transposable elements, antisense RNA and ribozymes seem afoot. Only gene transfers to cereals are still difficult and cumbersome. But this seems to be only a question of time and before long it is expected that even transformation *in vivo* would be possible which is when cumbersome procedure of tissue culture step will not longer be required. Also, the DNA transfer are not targetted giving rise to variation in the transformants for a particular character. This variation is explained by assuming a position effect. Altogether, there is still much more beyond that is challenging and should keep us busy for a while.

## References

- 1 Barany F, Kahn M E and Smith H O 1983 Directional transport and integration of donor DNA in *Haemophilus influenzae* transformation, *Proc Natl Acad Sci USA*, **80** 7274-7278
- 2 Bernal J D 1969 *Science in History* (Harmondsworth) Fox M S and Allen M K 1964 On the mechanism of deoxyribonucleate integration in pneumococcal transformation; *Proc Natl Acad Sci USA* **52** 412-416  
— — 1966 On the mechanism of integration of transforming deoxyribonucleate, *J Gen Physiol* **49** 183-196
- 3 Hershey A 1966 The into cells by phage, in *Phage and the Origins of Molecular Biology* pp 100-108 eds J Cairns, G Stent and J Watson, (Cold Spring Harbor)
4. Joshi V P, Kanade R P and Notani N K 1984 Genetic transformation with chimeric plasmid DNA in *Haemophilus influenzae* in *Proc 15th Intl Congress of Genetics* Vol I 167-171 eds V L Chopra, B C Joshi, R P Sharma and H C Bansal, (N Delhi)  
— — and Notani N K 1988 Decreased extrachromosomal fixation of chimaeric plasmid in strain N19 of *Haemophilus influenzae* Rd; *J Genet* **67** 113-119
5. Kahn M, Maul G and Goodgal S H 1982 Possible mechanism for donor DNA binding and transport in *Haemophilus*, *Proc. Natl. Acad Sci USA* **79** 6370-6374  
— — Barany F and Smith H O 1983 Transformasomes: specialized membranous structures which protect DNA during *Haemophilus* transformation; *Proc Natl Acad. Sci USA* **80** 6927-6931
6. Kanade R P and Notani N K 1987 Cloning and characterization of DNA repair gene in *Haemophilus influenzae*, *J. Biosci.* **12** 115-123
7. Lacks S 1962 Molecular fate of DNA in genetic transformation of pneumococcus *J Mol. Biol.* **5** 119-131

8. Medawar P 1984 *The Limits of Science*, Oxford
9. Mody R, Joshi V P and Notani N K 1990 in *Proc. Symp. Environmental Mutagenesis and Carcinogenesis* Vol I 1-6, ed P S Chauhan, Bombay
10. Notani N K and Goodgal S H 1966 On the nature of recombinants formed during transformation in *Haemophilus influenzae*; *J. Gen. Physiol.* **49** 197-209
 

— — 1971 Genetic and physical properties of unintegrated donor DNA molecules during *Haemophilus* transformation; *J. Mol. Biol.* **59** 223-226

— —, Joshi V R and Gopal-Ayengar A R 1971 Basis for radiosensitivity of some mutants of *Haemophilus influenzae*; in *Proc. Symp. Radiation and Radioisotopes for Industrial Microorganisms* 43-51, Vienna

— — — Setlow J K, Joshi V R and Allison D 1972 Molecular basis for the transformation defects in mutants of *Haemophilus influenzae*, *J. Bacteriol.* **110** 1171-1180

— — — — McCarthy D and Clayton N L 1981 Transformation of *Haemophilus influenzae* by plasmid RSF 0885 *J. Bacteriol.* **148** 812-816

— — — 1981 Genetic and molecular events in transformation of *Haemophilus influenzae* with plasmid RSF 0885 carrying cloned segments of chromosomal DNA; *J. Biosci.* **3** 431-438

— — —, Notani N K, McCarthy D and Clayton N L 1981 Transformation of *Haemophilus influenzae* by plasmid RSF 0885 containing cloned segments of chromosomal DNA, *J. Bacteriol.* **148** 804-811
11. Stent G S 1969 *The Coming of the Golden Age a view of the end of progress*, New York
 

— — 1978 Prematurity and uniqueness in scientific discovery in *Paradoxes of Progress* 95-113, San Francisco
12. Tuli R, Saluja J and Notani N K 1989 Cloning and expression in *Escherichia coli* of entomotoxic protein gene from *Bacillus thuringiensis* subspecies *Kurstaki*, *J. Genet.* **68** 147-160
13. Viegas P M, Mathews H, Bhatia C R and Notani N K 1987 Monohybrid and dihybrid segregations in the progenies of tobacco transformed for kanamycin resistance with Ti-vector system; *J. Genet.* **66** 25-31
 

— — and Notani N K 1988 Transgenic plants: Horizontal transfer of *nif* HD DNA to tobacco, *Indian J. Biochem. Biophys.* **25** 523-527

— — (unpublished)

— — and Notani N K 1990 Early fate in dry cereal embryos/seeds exposed to *H. influenzae* DNA; *Proc. D A.E. Symp. on Advances in Molecular Biology* 476-478



**Pushpa Mittra Bhargava** (b. 22 February 1928) did Ph D. (1950) from Lucknow University and was awarded D.Sc. (h.c.) of University of Burdwan. He was CSIR Distinguished Fellow (1990-92); Director, (till 1990), Centre for Cellular and Molecular Biology (CCMB), Hyderabad.

Bhargava discovered multifunctional, antifertility and anti-HIV protein, seminal-plasmin, and its possible role in regulation of fertility and HIV infection. Developed techniques of making liver cell suspensions and established their many uses.

Demonstrated that injected liver cells home to liver. Studied the structural and chemical composition of liver. Developed a model based on permeability controls to explain malignant transformation and normal cell division. Worked on the 'second genetic code' (in tRNA). Made findings that relate to the origin of life, specially the transition from chemical to biological evolution (formation of the first cell). Demonstrated the autonomy of mitochondrial transcription and translation. He has co-authored 'Proteins of Seminal Plasma' (John Wiley, 1989), edited 'Nucleic Acids: Structure, Biosynthesis and Function' (CSIR, 1965).

Bhargava is Fellow of Indian Academy of Sciences, National Academy of Sciences (India), National Academy of Medical Sciences (India), Andhra Pradesh Academy of Sciences, World Academy of Arts and Sciences, and New York Academy of Sciences; Corresponding Member, American Association for Cancer Research; and of several other international/national professional societies. He is the recipient of Watumull Memorial Prize (1962); FICCI Award (1979); B.N. Chopra Memorial Lecture Award (INSA) (1989); Ranbaxy Research Award (1990); SICO Award (1991); Goyal Prize (1993), Visiting Fellow and Life Fellow, Clare Hall, University of Cambridge, UK (1983, 1987); Distinguished Scientist of Andhra Pradesh (Andhra Pradesh Academy of Sciences) (1988); Golden Jubilee Scroll of Honour (Indian Society of Genetics and Plant Breeding) (1991); Padma Bhushan (1986).

*Pushpa Mittra Bhargava was elected to the fellowship of the Academy in 1979.*

**Address:** CSIR Distinguished Fellow, Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007; Tel: (0842) 854148; Telex: 0425-7046 CCMB IN; Fax: 91-842- 851195; Res: C-3, Regional Research Laboratory Houses, Hyderabad 500 007; Tel: (0842) 853498.

## SEMINALPLASMIN, FERTILITY AND AIDS

PUSHPA M BHARGAVA

Seminalplasmin, a 47-aminoacid-long protein discovered in our laboratory, is a potent inhibitor of bacterial growth. At high concentrations, it is bacteriocidal, and at even higher concentrations it lyses bacterial cells; lysis-resistant mutants of *E.coli* have been obtained and mapped by transduction. It also lyses certain mammalian cells. It kills a wide variety of microbes — gram-positive and gram-negative bacteria and yeasts—by entering the cell and inhibiting RNA polymerases, for examples, from *E. coli* and yeast, and reverse transcriptases, by binding strongly to the enzyme. It binds to DNA and unwinds it. Seminalplasmin, which is secreted by the accessory sex glands of bull but not by testes or epididymis, also influences spermatozoal functions. It binds to the plasma and acrosomal membranes of the spermatozoa and, as a consequence, brings about an increase in the fluidity of the membranes. Such changes are known to influence the acrosome reaction and fertilising ability of seminalplasmin. In fact, seminalplasmin inhibits the motility, acrosome reaction and fertilising ability of mammalian spermatozoa *in vitro*. In rats, it also inhibits *in vivo* fertilisation. Seminalplasmin also binds to DNA strongly and unwinds it.

Antiseminalplasmin, another protein purified to homogeneity from bull seminal plasma and  $\text{Ca}^{2+}$ , reverses all the biological effect of seminalplasmin on cells. Hence, it is possible that the effects of seminalplasmin on spermatozoa are related to  $\text{Ca}^{2+}$ , since  $\text{Ca}^{2+}$  is known to be required for motility, acrosome reaction and fertilising ability of spermatozoa. Seminalplasmin does not bind to  $\text{Ca}^{2+}$  but is capable of inhibiting the uptake of  $\text{Ca}^{2+}$  in spermatozoa. It appears to do so by binding to  $\text{Ca}^{2+}$  binding sites on spermatozoal membranes and thus reducing the amount of  $\text{Ca}^{2+}$  that is translocated across the membrane. Seminalplasmin is, in fact, a potent antagonist of calmodulin, to which protein it binds much better in the presence of  $\text{Ca}^{2+}$  than in its absence. The above observations put together indicate that the fertilising capacity of spermatozoa may depend on the relative concentration of seminalplasmin, antiseminalplasmin and  $\text{Ca}^{2+}$  in a given semen sample. Experiments are in progress to evaluate the hypothesis that the low rates of fertility sometimes



observed after artificial insemination in cattle are due to high levels of seminalplasmin in seminal plasma.

Seminalplasmin also inhibits the binding of anti-CD4 (but not anti CD2 or anti CD8) antibodies to CD4-positive lymphocytes and infection of lectin-stimulated human lymphocytes by HIV; it can, in fact, displace HIV bound to the surface of lymphocytes. These observations suggest that seminalplasmin may be useful as an anti-AIDS agent. Recent evidence indicates that a seminalplasmin-like protein is probably naturally present also on the surface of CD4-positive lymphocytes; its presence on these cells suggests that it may play a role in the prevention or/and development of the AIDS infection. Seminalplasmin has been chemically synthesised on a gram scale and the synthetic seminalplasmin found to be identical to natural seminalplasmin in all respects including its biological activities.

Since seminalplasmin is a small protein with many biological properties it could also be an ideal protein for studying structure-activity relationships following site-directed mutagenesis. With this in mind, the synthetic gene for seminalplasmin has been successfully cloned in *E coli* by two independent groups, i.e., at the Max Planck Institute for Biophysical Chemistry in Germany and at the Laboratory of Molecular Biology, Cambridge, UK. The gene for seminalplasmin has also been synthesised in the CCMB and is now being cloned in *E coli*

## References

1. Agustín J S, Hughes P and Lardy H 1987 Properties and function of caltrin, the calcium-transport inhibitor of bull seminal plasma; *FASEB J.* **1** 60-66
2. Aumüller G and Scheit K H 1987 Immunohistochemistry of secretory proteins in the bull seminal vesicle; *J Anat.* **139** 43-48
3. Bhargava P M 1981 Four new and unusual proteins from bovine seminal plasma; *Biochemical Society Transactions* **9** 540-545  
— 1981 The romance of the seminal fluid. The excitement of the unusual. Presidential address, Biology Section, National Academy of Sciences (India), in *Proceedings of the Academy, 1981, Annual number*, pp 51-67  
— 1986 Seminalplasmin, an antimicrobial and transcription inhibitory protein from bovine seminal plasma, might be nature's own antifertility agent; *Annals Natl Acad. Med. Sci (India)* **21** 51-68  
—, Jamil K, Rao N.S, Murthy B S N and Ramaswamy V 1986 Variation in the antibacterial activity of bovine seminal plasma; *Indian Vet. J.* **63** 642-649

- , Reddy E S P, Rao V N and Scheit K H 1980 Seminalplasmin, a new protein inhibitor of microbial growth and transcription from bovine seminal plasma, and two other proteins of seminal plasma; in *Biomolecular Structure, Conformation and Function* ed. R. Srinivasan, Pergamon Press (London), Vol. 2, pp. 177-189
4. Chitnis S, Prasad K S N and Bhargava P M 1987 Bacteriolytic activity of seminalplasmin; *J. Gen. Microbiol.* **133** 1265-1271
  5. Comte M, Malroe A and Cox J A 1986 *Biochem. J.* **240** 567-573
  6. Galla H J, Warncke M and Scheit K H 1985 Incorporation of the antimicrobial protein seminalplasmin into lipid bilayer membranes; *Eur. Biophys. J.* **12** 211-216
  7. Gietzen K and Galla H J 1985 Seminalplasmin: an endogenous calmodulin antagonist; *Biochem. J.* **230** 277-280
  8. Gopal V, Prasannakumar K and Chatterji D 1988 Bovine seminalplasmin is a DNA unwinding protein; *FEBS Letters* **231** 389-393
  9. Heaphy S, Singh M and Gait M J 1987 Cloning and expression in *E. coli* of synthetic gene for the bacteriocidal protein caltrin/seminalplasmin; *Protein Engineering* **1** 425-431
  10. Kemme M, Madiraju-Murthy V V S, Krauh E, Zimmer M and Scheit K H 1986 The major protein of bull seminal plasma is a secretory product of seminal vesicle; *Biochimica et Biophysica Acta* **884** 282-290
  11. Lardy H and Agustun J S 1989 Caltrin and calcium regulation of sperm activity; in *The Cell Biology of Fertilisation*, pp. 29-39 (New York: Academic press)  
— and — 1989 Caltrin and calcium regulation of sperm activity; in *The Cell Biology of Fertilisation*, pp. 29-39 (New York: Academic Press)
  12. Levis R V, Agustin J S, Kruggel W and Lardy H A 1985 The structure of caltrin, the calcium transport inhibitor of bovine seminal plasma; *Proc. Natl. Acad. Sci. USA* **82** 6490-6491
  13. Pandit M W 1985 Prediction of strong antigenic determinant of seminalplasmin and ribonuclease from the amino acid sequence; *Int. J. Pep. Prot. Res.* **25** 197-199
  14. Preuss K D, Sachse H, Sharma H S, Hagendorff G and Scheit K H 1987 Molecular cloning and expression of a synthetic DNA coding for the antimicrobial protein of bull seminal plasma; *Biol. Chem. Hoppe-seyler* **368** 501-506
  15. Rao V N and Bhargava P M 1985 Isolation, characterisation and possible mode of action of antiseminalplasmin, a new protein that inhibits the antimicrobial activity of seminalplasmin; *Biochemical Journal* **227** 609-619  
—, Reddy E S P, Torriani A and Bhargava P M 1983 Seminalplasmin inhibits transcription and translation of DNA *in vitro*; *FEBS Letts.* **152** 6-10

16. Reddy E S P and Bhargava P M 1979 Seminalplasmin: an antimicrobial protein from bovine seminal plasma which acts in *E. coli*. by specific inhibition of rRNA synthesis; *Nature* **279** 725-728  
—, Das M R, Reddy E P and Bhargava P M 1983 Inhibition of reverse transcriptase by seminalplasmin; *Bioch. J.* **209** 183-188
17. Scheit K H 1986 The major basic proteins of bull seminal vesicle secretion; *Hoppe-Seyler's Z. Biol. Chem.* **367** 229-233  
— and Bhargava P M 1985 Effect of seminalplasmin, an antimicrobial protein from bovine semen, on growth and macromolecular synthesis in *Candida albicans*; *Indian J. Biochem. Biophys.* **22** 1-7  
—, Reddy E S P and Bhargava P M 1979 Seminalplasmin is a potent inhibitor of *E. coli* RNA polymerase *in vitro*; *Nature* **279** 728-731  
—, Shivaji S and Bhargava P M 1985 Seminalplasmin, an antimicrobial protein from bull seminal plasma, inhibits growth, and synthesis of nucleic acids and proteins in *S. cerevisiae*; *J. Biochem.* **97** 463-471  
— and Zimmer M 1984 Seminalplasmin, an antimicrobial protein from bull semen, inhibits gene expression in *E. coli*; *Biochim. Biophys. Acta* **781** 187-191  
—, Rao N S, Prasad K S N and Bhargava P M 1986 Antibacterial activity of seminalplasmin, a basic protein from bovine seminal plasma; *Microbios* **45** 105-112
18. Shivaji S 1984 Antimicrobial activity of semen; *Trends Biochem. Sci.* **8** 101-107  
— 1986 Binding of seminalplasmin to the plasma and acrosomal membranes of bovine spermatozoa. Fluorescence studies on the changes in the lipid-phase fluidity; *FEBS Lett.* **196** 255-258  
— 1988 Seminalplasmin: a protein with many biological properties; *Biosci. Rep.* **8** 609-617  
— 1987 Interaction of seminalplasmin with chlortetracycline, a fluorescent chelate probe  $\text{Ca}^{2+}$ ; *FEBS Letts* **218** 97-101  
— and Bhargava P M 1987 Antifertility factors of mammalian seminal fluid; *Bioassays* **7** 13-17  
—, — and Scheit K H 1984 Immunological identification of seminalplasmin in tissue extracts sex glands of bull; in *Protides of the Biological fluids* ed H Peeters, (Oxford and New York: Pergamon Press) A7 pp. 281-284  
—, Scheit K H Bhargava P M 1989 *Proteins of Seminal Plasma* (New York: John Wiley & Sons)
19. Sitaram N and Chatterji D 1985 Interaction of bovine seminalplasmin with *Escherichia coli* RNA polymerase in the presence of rifampicin; *FEBS Lett.* **182** 77-80

- , Kumari V K and Bhargava P M 1986 Seminalplasmin and caltrin are the same protein; *FEBS Lett* **201** 233-236
20. Theil R and Scheit K H 1983 Amino acid sequence of seminalplasmin, an antimicrobial protein from bull semen; *EMBO J* 1159-1163
- and Scheit K H 1983 Purification and spectral characterisation of seminalplasmin, an antimicrobial protein from bull semen; *Hoppe Seyler's Z Physiol Chem* **361** 1003-1009
21. Venkov P and Scheit K H 1984 Effects of seminalplasmin on rRNA synthesis in *Saccharomyces cerevisiae*; *FEBS Lett* **172** 21-24



**Thamarapu Vedantha Desikachary** (b. September 1919) did D.Sc. (1964) from University of Madras. He retired as Professor of Botany, University of Madras, Madras.

Desikachary's research achievements are in algae (cyanophyta, marine plants, taxonomy and biology of blue- green algae) and Volvocales (with M.O.P.Iyengar). Along with B.G.L Swamy, he propounded a theory of isolation and insulation of special cells, and, with Sundaralingam, a new concept on the interrelationship of Charophyta.

Desikachary is Fellow of Indian Academy of Sciences and Phycological Society of India. He is the recipient of Professor C. V. Raman Medal (University of Madras); Professor V. Puri Medal (Indian Botanical Society); Professor Panchanan Maheshwari Memorial Lecture Award (INSA) (1987)

*Thamarapu Vedantha Desikachary was elected to the fellowship of the Academy in 1966*

**Address:** Emeritus Professor, Old Botany Lab. Buildings, University of Madras, Madras 600 005; Res 121, Santhome High Road, Madras 600 028; Tel: (044) 75047

## **SOME THOUGHTS ON THE CLASSIFICATION OF RED AND GREEN ALGAE**

**T V DESIKACHARY FNA**

I consider it a distinct and great honour to have been invited by INSA to deliver this Memorial Lecture. The name of Professor Maheshwari is a name to be conjured with. He was an eminent morphologist, largely built by his own endeavours, characterized by an unparalleled zeal, and propelled by his devotion to the subject only to become an erudite and a respected botanist. I have had the privilege of coming closer to him than my position would warrant. I was happy that the Professor developed an affection for and trust in me. That was because I was Professor Iyengar's student. I am very grateful indeed for this honour.

I shall in the short time at my disposal refer to some aspects of the morphology and development of algae and their bearing on taxonomy, especially, of the red algae and the green algae.

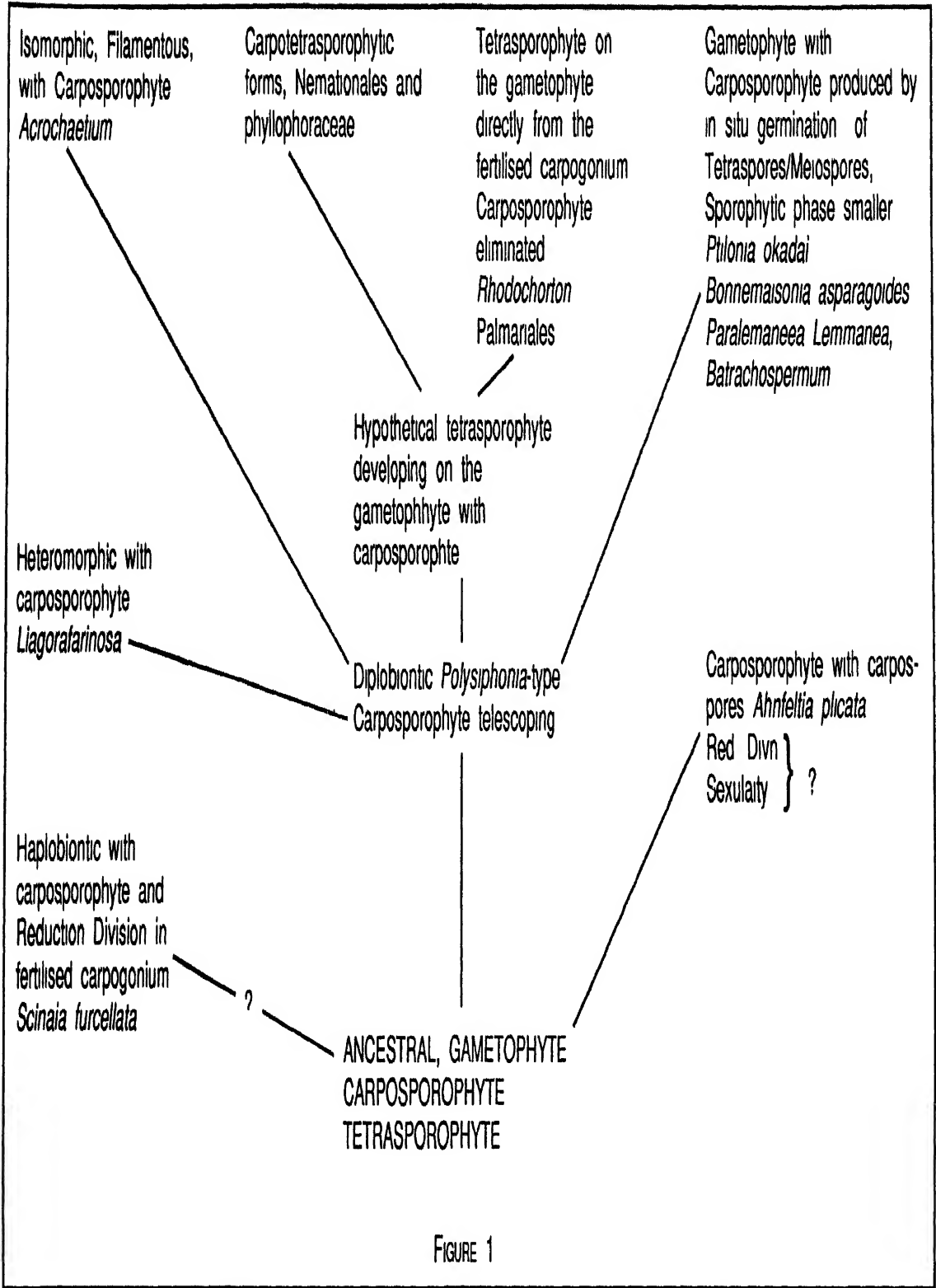
The red algae do not have motile forms or flagellated stages to draw upon one's experiences with Nature's experiments in evolution, to build a system arising from such simple to complex forms. Elsewhere, in some algae, with the development of sexuality one observes the inception of reduction division in the fertilized zygote before germination or further development, and any postponement of meiosis naturally leads to a diplobiontic life cycle with a gametophyte and a sporophyte. This diplobiontic life cycle has generated many theories to develop on its origin. In Phaeophyceae where we do not have many simple forms we are more or less stuck with a predominant support for the Homologous theory. We have a different situation in the Rhodophyceae where the Antithetic theory held sway for a long time until Feldmann came out with a similar support for the Homologous theory even in the case of the red algae (Feldmann 1972).

Among extant red algae fertilization of the carpogonium by a spermatium sets in motion a more or less unique system of asexual multiplication by diploid mitospores (carpospores) formed from the fertilized carpogonium. The entire structure may be compact (cystocarp) with a multicellular wall like structure formed by the overgrowth of gametophytic tissue. Or it may be a loose bouquet of branches, with sterile

filaments, involucre filaments, intermixed with the fertile branches. The fertile portion in both the cases is sporophytic, the carposporophyte, producing carpospores. For a long time, reduction division was presumed to take place, in many members of the Nemalionales, in the fertilized carpogonium as against the more common occurrence of it in a free living sporophyte (the tetrasporophyte) in sporangia producing tetraspores (or bispores, rarely polyspores). Phycologists put forth many theories largely repetitive of the two major concepts, modifications of the homologous and Antithetic theories. Many of the life history types are considered as derived ones with one or more phases in the typical diplobiontic theory laid stress on forms with tetraspores formed, presumably after reduction division, in carposporangia instead of single carpospores being formed in each carposporangium. Supporters of homologous theory depended upon the *Polysiphonia* — type and the problematic monospore producing *Ahnfeltia plicata* to suggest that the hypothetical ancestor has had two sporophyte phases, one producing mitospores or monospores (carpospores?) and the other meiosporic tetraspores. Telescoping or reduction or elimination of one or more phases and their functions being hoisted on to the other phases has led to myriad varieties of life history patterns seen in the red algae. It must be conceded that the red algae exhibit varied patterns of life histories that easily suggest elaboration and or abridgement of one or more facets or phases in the vegetative phase and asexual reproduction. Another complicating factor is the absence (or nondiscovery) of sexuality in a number of forms (figure 1).

Antithetic theory depends on a linking up of phenomena observed among the extant forms and a meaningful assertion that a step by step postponement of reduction division and a consequent progressive elaboration of the diploid vegetative phase leads to an alternation of (isomorphic or heteromorphic) generations of a gametophyte and an independent sporophyte. One immediate consequence of the postponement or a shift of the location of the reduction division from the fertilized carpogonium is the development of the carposporophyte with diploid carpospores (figure 2).

Desikachary (1957) dealing with members of the section *Mucosae* of *Liagora* suggested that should the diploid carposporophyte be confirmed as giving rise to rhizoids, as has been described by earlier workers then such a carposporophyte may represent an attempt by the carposporophyte to become nutritionally independent of the gametophyte. Such carposporophytes have been described in *Trichogloeopsis* by Doty





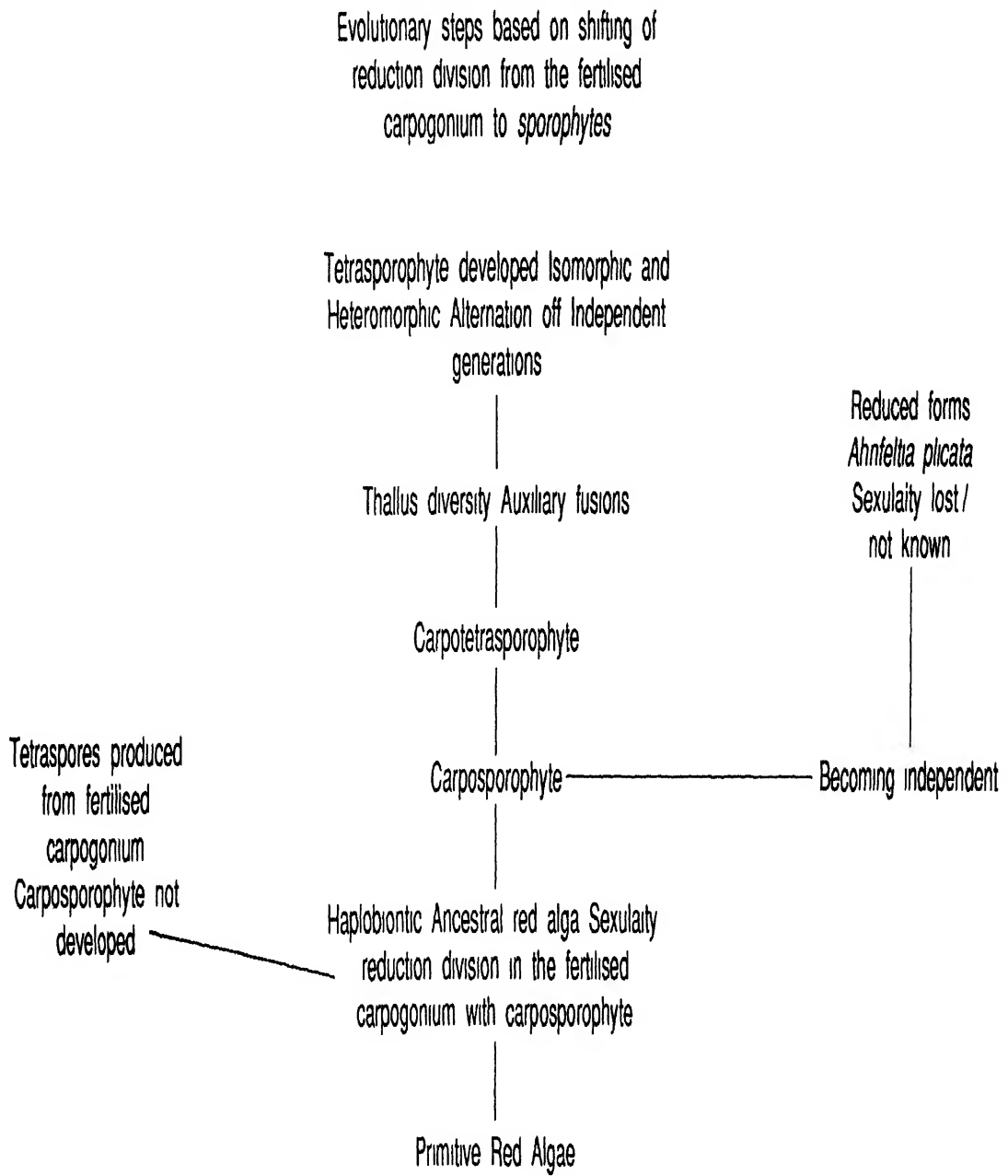


FIGURE 2

and Abbott. One does not know whether they also serve as aids in nutrition. However, in most red algae the carposporophytes have remained anchored to and dependent or parasitic on the gametophyte.

Red algal thallus has no great variation in construction or tissue differentiation and hence it has very little to offer by way of provoking thoughts on evolution. Naturally all the work of the last century largely centered round the unravelling of the ontogeny of the carposporophyte.

The developing zygote or fertilized carpogonium exhibits many features in what are basically only nutritive fusions:

(1) *By direct access:*

Enlargement of the connections between the sterile cells of the carpogonial branch and the supporting cells; the unmodified part product of the fertilized carpogonium sitting on the top of large fusion cell at the junction with the carposporophyte.

(ii) *Auxiliary cell fusions negotiated through special connections:*

- (a) with special cells of the carpogonial branch and the supporting cells, by hypha like cells or connecting filaments or cells.
- (b) with other homologous cells but situated on special branches, by connecting filaments,
- (c) with distinguishable vegetative cells in the neighbourhood by a connecting filament,
- (d) with special cells formed for this purpose by the supporting cells, and
- (e) random fusions with cells (Nurse cells) in the neighbourhood of the carpogonium.

Such fusions do not occur in the Acrochaetiales where the carpogonial branches are sessile.

Carpospore formation may be restricted to the fertilized carpogonium or the fusion sites (cells) or may be indeterminate in that they may arise from different parts of the haustorial filament or in the vicinity of sites of the fusions.

It is long recognized that these fusions do not participate actively in the reproductive processes as such but are largely nutritive in function and

at best form centres of spore production with the transfer of the products of the fertilized nucleus to it (Desikachary 1982). The red algae are classified into Orders on the basis of post-fertilization changes and these Orders are generally accepted. Differences lie in newer understanding of the Nemalionales (sensu lat.)

The advent of cultures has been an eye opener and it has yielded many unknown or cryptic phases of life history. This has been responsible for a rethinking on many accepted 'facts'. The impact of a perennial aquatic environment seems to have off-set the need for a perennating or a resting phase in the life history. It has been demonstrated that many plants have an encrusting filamentous or discoid juvenile stage ('adelophycean' stages) from which the more obvious phases develop during appropriate seasons. Thus diplobiontic rather than haplobiontic plants are the rule among the red algae. This situation leaves fewer simpler forms for us to play with in developing ideas on evolution.

The basic phenomenon seen in many algae is the occurrence of reduction division in the fertilized zygote. The red algae do not have many such instances. In the Bangiophycidean algae, in cases with 'alleged' sexual reproduction, the fertilized carpogonium becomes totally converted into a sporangium and the alternate phase has monospores (conchospores) with reduction division yet to be located or demonstrated. The nearest to such a condition is seen among the red algae in *Rhodophysema* in which the fertilized carpogonium cuts off a sterile cell and becomes converted into tetrasporangium. From the fertilized carpogonium to imagine the development of a loosely branched (or a nemathecial) carposporophyte does not need any great interpolation. Similarly, the occurrence of tetraspores in a carposporophyte and further a development of a tetrasporophyte is seen in one and the same genus (as at present conceived) of the Nemalionales. So while there is a great deal of evidence for the totally predominant diplobiontic basic set up to the near elimination of the haplobiontic make up, yet the presence of these instances and the complexities seen in the Nemalionales suggest strongly that the haplobiontic condition might still be ancestral.

Recent studies on *Rhodophysema* have shown that its life history approximates to the classical haplobiontic life cycle so characteristic of primitive algae, production of meiospores from the fertilized carpogonium without forming a large thallus.

Among the species of *Rhodochorton* and allied genera we have, significantly, both instances of formation of carposporophytes and their suppression. We are interested here in the latter ones. We have a direct development of the tetrasporophyte from the gonimoblast cells. Here again there is successive production of spores from a stalk like cell cut off by a transverse division of the fertilized carpogonium. It is these cells or spores that develop into tetrasporophytes which developing rhizoids become independent (Feldmann 1972, West 1969, Magne 1970, Steganga 1978, Ohta & Kurogi 1979).

*Halosaccion ramentaceum* has a diphasic life history and lacks carposporophytic stages. The carpogonium is a simple cell and there are no auxiliary cells or fusions. The fertilized carpogonium develops directly into a tetrasporophyte which becomes independent by developing a hold fast (Van der Meer 1981).

In *Rhodophysema*, the fertilized carpogonium enlarges and functions directly as a sporangium. The contents divide into two unequal cells, an upper tetrasporangial cell and a lower stalk cell which can form successive generations of tetrasporangia by a repetition of similar unequal divisions. Sometimes the tetrasporangia germinate *in situ* and produce gametophytes. It is possible *Ahrfeltia plicata* may be found to be similarly lacking a carposporophyte but reproducing similarly sexually (Gregory 1935). In *Palmaria palmata* also carposporophytic stages are absent. The tetrasporophyte develops from the fertilized carpogonium and becomes independent by developing diploid rhizoids later than sooner (Van der Meer & Todd 1980 Decen & West 1982).

We have similar instances of germination of meiospores in the creeping filamentous ('Chantransia') stages of some Nemalionales, *Batrachospermum*, *Lemnaea* etc., and there is no discharge or dispersal of the spores, a tendency we have already seen in the Palmariales.

Thus we have a number of algae which can usefully illustrate the gradual shifting of the reduction division from the fertilized zygote producing meiospores leading to the formation of a diploid carposporophyte and then to an independent tetrasporophyte leading to a derivation of *Polysiphonia*-type of individuals. These would then involve: i) that the primitive red alga would be the one which has reduction division taking place in the fertilized zygote and ii) an extensive evolution of complex systems of fusions to promote nutrition and multiplication of the individual by spore production by diploid carposporophyte. All other

types may have to be thought off as derived from this archetype by a variety of types of sporophytic phases by a postponement of reduction division and attempts by the sporophyte to achieve independence. Homologous theory would involve a suppression of many phases which is probably a negation of evolution with such a great deal of expression seen in present day forms would allow us to accept or presume. The classification recommended here is largely based on the development of the carposporophyte and the fusion patterns.

A synoptic classification of the red algae is presented here.

BANGIOPHYCIDAE

PORPHYRIDIALES

Porphyridiaceae

GONIOTRICHALES

Goniotrichaceae

Phragmonemataceae

BANGIALES

Bangiaceae

ERYTHROPELTIDALES

Erythropeltidaceae

Smithoraceae

Boldiaceae

Rhodochaetaceae

Compsopogonaceae

ACROCHAETIALES

Acrochaetiaceae

PALMARIALES

Palmariaceae

BATRACHOSPERMALES

Batrachospermaceae

Lemaneaceae

Thoreaceae

NEMALIONALES

Nemaliaceae

Trichogloeopsidaceae

Helminthocladiaceae

Dermonemataceae

**CHAETANGIALES**

Chaetangiaceae

**BONNEMAISONIALES**

Bonnemaisoniaceae

Neccaraceae

**GELIDIALES**

Gelidiaceae

Gelidiellaceae

**CRYPTONEMIALES**

Halymeniaceae, Peyssonneliaceae, Kallymeniaceae, Dumontiaceae,  
 Weeksiaceae, Corynomorphaceae, Gloiosiphonaceae,  
 Petrocladiophyllaceae, Choreocolacaceae, Pseudoanemoniaceae,  
 Endocladiaceae, Trichocarpaceae

**CORALLINALES**

Corallinaceae

**HILDENBRANDIALES**

Hildenbrandiaceae

**GIGARTINALES**

Gigartinaceae, Gracilariaceae, Phyllophoraceae, Solieriaceae,  
 Gymnophlaeaceae, Calosiphoniaceae, Blinksiaceae, Caulacanthaceae,  
 Phacelocarpaceae, Nizymeniaceae, Cystocloniaceae, Acrotylaceae,  
 Dicranemataceae, Mychodeaceae, Mychodeophyllaceae, Plocamiaceae,  
 Sebdeniaceae, Furcellariaceae, Sarcodiaceae, Rissoellaceae,  
 Sphaerococcaceae, Chondriellaceae, Cruoriaceae, Hypneaceae,  
 Polyidaceae, Rhizophyllidaceae, Cubiculosporaceae.

**RHODYMENIALES**

Rhodymeniaceae, Champiaceae, Lomentariaceae

**CERAMIALES**

Ceramiaceae, Delesseriaceae, Dasyaceae, Rhodomelaceae

One word on the Order Palmariales. The Order Palmariales is unique in one feature, i.e., lack of a carposporophyte. We are unable to comprehend that the presence of a stalk cell is of any greater importance as similar first divisions do occur in a number of red algal genera. In fact this type of division into an active cell and latent (stalk) cell is common in many divisions preceding reproduction, asexual or sexual. The main

feature is the absence of a carposporophyte. The second important feature is the lack of any post-fertilization fusion.

## GREEN ALGAE

In the green algae, considerations of the evolution of the higher plants from the algae and the presence of many haplobiontic forms, led people to accept concepts similar to the antithetic theory. But all is not well and exercises in theories are not as common as in the reds and browns. In these again, life history of marine forms and EM studies have brought in a little bit of shock and we are in a disturbed state. The following is a synoptical outline of the classification of these.

### PYRAMIDOMONADALES

Pedinomonadaceae, Mantoniellaceae, Pyramimonadaceae,  
Nephroselmidaceae, Platymonadaceae

### VOLVOCALES

Dunaliellaceae, Raciborskiellaceae, Oltmannsiellaceae,  
Chlamydomonadaceae, Phacotaceae, Haematococcaceae,  
Basichlamydeaceae, Spondylomoraceae, Astrephomenaceae, Volvocaceae.

### TETRASPORALES

Gloeococcaceae, Chlorangiaceae, Tetrasporaceae, Chaetopeltidaceae.

### CHLOROCOCCALES

Palmellaceae, Palmellopsidaceae\*, Chlorococcaceae, Characiaceae,  
Chlorellaceae, Oocystaceae, Chlorodendraceae, Dictyosphaeriaceae,  
Coelastraceae, Botryococcaceae, Scenedesmaceae, Protosiphonaceae,  
Characiosiphonaceae, Hydrodictyaceae, Hydrodictyopsidaceae\*,  
Eremosphaeraceae, Gomontiaceae.

### ULOTRICHALES

Ulotrichaceae, Microsporaceae, Cylindrocapsaceae,  
Cylindrocapsopsidaceae\*, Chaetosiphonaceae, Microthamniaceae.

### KLEBSHORMIDIALES:

Klebshormidiaceae

### CHAETOPHORALES

Chaetophoraceae, Wittrockiellaceae, Jaoaceae, Chlogrosphaeraceae,  
Aphanochaetaceae, Chaetophaeridiaceae

### COLEOCHAETALES:

Coleochaetaceae

**CHROOLEPIDALES:**

Chroolepidaceae (Trentepohliaceae)

**OEDOGONIALES:** Oedogoniaceae**ULVALES**

Ulvaceae, Monostromaceae, Ulvopsidaceae, Capsosiphonaceae

**SCHIZOMERIDALES:** Schizomeridaceae**PRASIOALES:** Prasiolaceae**ZYGNEMATALES**

Mesotaeniaceae, Zygnemataceae, Mougeotiaceae, Gonatozygaceae

**DESMIDIALES :** Desmidiaceae**CLADOPHORALES:** Cladophoraceae, Arnoldiellaceae, Acrosiphoniaceae, Anadyomenaceae**SIPHONOCLADALES:** Siphonocladaceae, Boodleaceae**SPHAEROPLEALES:** Sphaeropleaceae**DICHOTOMOSIPHONALES:** Dichotomosiphonaceae, Phyllosiphonaceae**DERBESIALES:** Derbesiaceae**CODIALES:** Codiaceae, Bryospsidaceae**CAULERPALES:** Caulerpaceae**VALONIALES\*:** Valoniaceae**DASYCLADALES:** Dasycladaceae, Acetabulariaceae**CHARALES**

Characeae, Eocharaceae, Palaeocharaceae, Clavatoraceae, Trochiliscaceae, Sycidiaceae

The synoptic classification presented here includes Orders and Families generally accepted by phycologists. Some of these have yet to be validated/conserved\*. It is characterized by many personal opinions and preferences. A detailed analysis of characters and constituents must await some more time. The Prasinophyceae, it is felt, is yet to be finally shaped as the number of species worked out are few. Generitypes have not been studied in quite a few cases. Unstudied species are large and therefore are yet to be taxonomically reallocated. There is a hesitancy to use EM characters in the recognition of taxa except when they are linked with characters observable under the light microscope. More and more



characters are being revealed necessitating modifying our opinions as in the case of *Mantoniella*. The practice of dividing the green algae into a number of classes, as was first suggested by Pascher and later not fully followed even by him, is not followed as such a practice is not realistic and overshadows the intricate and complex affinities between the different Orders. We hold the concept that without protoplasmic connections there is no true parenchymatous thallus as very relevant among on habit, external form and pattern of reproduction and life history. The import of heterotrichy is for the present accepted but one feels that it may have to reinvestigated as many 'Ulotrichales' have shown, in culture, prostrate systems. Stephanokonton conditions are more prevalent than earlier realized and basing of classes on this character alone is not considered a valid one except in conjunction with other exclusive characters. Ontogeny of branching is distinctive of the Cladophorales and Chroolepidales. Sporangiate forms are recognized as distinct. Segregative cell division is accepted as a very distinct feature and the Order Siphonocladales is recognized to the exclusion of the Cladophorales (Fritsch 1947). The Valoniales are considered distinct from the Siphonocladales as during division the entire protoplast of a cell does not take part as in the latter and only a mere lens shaped cell is cut off. In some cases such proliferations take place from the rhizoid rather than from the emergent portion of the thallus. Lastly that each of the siblings inherit one half of the parent's cell wall (as in the case of Diatoms) makes one feel that the Desmidiaceae (Placoderms) must be kept distinct.

I do realize that I have not been able to do proper justice to this extensive subject but I have covered as much as I can within the time at my disposal. I must thank everyone, —Professor H D Kumar, colleagues in the Centre of Advanced Study in Botany. BHU Professor Ramachandra Rao. Faculty Members of the BHU, and Fellow phycologists and Fellows of the INSA for their kindness and for the courtesies extended to me.

Thanking you

## References

1. Abbott, C A and Doty, M S 1960 Studies in the *Helminthocladiaceae* II *Trichogloeopsis*; *Ann J Bot* **47** 632-640
2. Decew, T C and West, J A 1982 A sexual life history of *Rhodophysemma* (Rhodophyceae) a reinterpretation; *Phycologia* **21** 67-74

3. Desikachary, T V 1982 On the nature of auxiliary cells in red algae; *Seaw Res Util* 5 49-52
4. Desikachary, T V 1957 On a "*Liagoropsis*" from Andamans; *Rev Algol (n s)* 3 51-56
5. Feldmann, J 1972 Les problèmes actuels de l'alternance de générations chez les algues; *Mem Soc Bot Fr* 7-38
6. Fritsch, F E 1946 The status of Siphonocladales; *Indian bot Soc M O P Iyengar Comm Vol* pp 29-48
7. Gregory, B D 1934 On the life — history of *Gymnogongrus griffithsiae* Mart and *Ahnfeltia plicata* Fries; *J Linn Soc Lond* 42 531-551
8. Magne, F M 1970 L'évolution du cycle de développement chez les Rhodophycées; *C R Acad Sci D* 271 1672-1674
9. Ohta, M. & Kurogi, M 1979 On the life history of *Rhodochorton purpureum* (Lightf) Rosenvinge from Hokkaido in culture; *Jap J Phycol* 27 161-167
10. Stegenga, H 1978 The life histories of *Rhodochorton purpureum* and *Rhodochorton floridulum* (Rhodophyta, Nemaliales); *Br phycol J* 13 279-289
11. Van der Meer, J P 1981 The life-history of *Halosaccion ramantaceum*; *Can J Bot* 59 433-436
12. —, J P & Todd, E R 1980 The life-history of *Palmaria palmata* in culture A new type for the Rhodophyta; *Can J Bot* 58 1250-1256
13. West, J 1969 The life histories of *Rhodochorton purpureum* and *R tenue* in culture; *J Phycol* 5 12-21



**Vishwambhar Puri** (b. 10 December 1909) did D.Sc. (1940) from Agra University. He is Professor Emeritus (1975- ), and was formerly Hony Director, Institute of Advanced Studies (1974-75), Senior Professor and Head, Botany Department (1969-72), at Meerut University.

Puri studied floral anatomy of Rhoeadales and Parietales; offered a new interpretation of crucifer gynoecium and clarified placentation in Cucurbitaceae and other families; offered a new interpretation of *Marsilea* sporocarp; and discussed the morphology of inferior ovary. Published two

exhaustive reviews in *Botanical Reviews*. His more recent studies on evolution have suggested that monocots and dicots are more independent than is usually believed. He has Prepared 'History of Botany (Floral Anatomy) for India, Burma and Ceylon', at the invitation of Indian Botanical Society.

Puri is Fellow of Indian Academy of Sciences, National Academy of Sciences (India), and Indian Botanical Society; has been Member of INSA Council (1977-79). He is the recipient of Birbal Sahni Medal (1964), K N. Bahl Medal (1988), Elected Hony Vice-President, XII International Botanical Congress, Leningrad (1975); Professor Panchanan Maheshwari Memorial Lecture Award (INSA) (1990)

The Indian Botanical Society has instituted a medal in his honour, awarded annually since 1981.

*Vishwambhar Puri was elected to the fellowship of the Academy in 1957*

**Address:** Emeritus Professor, Botany Department, Meerut University, Meerut;  
**Tel:** (0121) 75021/19; **Res:** 13, Tilak Road, Meerut 250 001; **Tel** (0121) 543455.

## **MONOCOTYLEDONS : SOME COMMENTS ON THEIR MORPHOLOGY AND EVOLUTION**

V PURI

Angiosperms, the contemporary of mammals, are the most recent of all land plants, being about 120 million years old, and yet they are the most numerous, widespread and varied in form, habit and habitat. These are usually distinguished into monocots and dicots, depending upon the number of cotyledons in the seed. The former are the smaller of the two groups and have been placed in classification sometimes before, sometimes after the dicots. This smacks of some doubt as to their status and importance. It may, therefore, be worthwhile to examine if this group that feeds the entire living world, hides the nakedness of mother earth, beautifies the landscape with orchids, lilies, strelitzias, etc., and decorates the tropical sea coast all over the world with majestic coconut palms, is having a fair deal at the hands of the taxonomist. And also to have a look at its fossil history to determine, if possible, the causes of its apparent heterogeneity.

### **TERMINOLOGY AND CLASSIFICATION**

Much of our botanical terminology is old and has become obsolete in some respects, based as it is in many cases, on characters whose concepts have undergone marked modifications as a result of enormous amount of subsequent work and consequent better understanding of their scope and extent. In the 18th century when de Jussieu proposed the terms monocotyledons and dicotyledons, these may have been all right, for not much was known about the exceptional situations. Besides, our information about other features was very scanty and meagre. But now the conditions have changed and the old terms appear rather inadequate.

Reacting to this situation, Cronquist, Takhtajan and Zimmermann (1966) suggested type-based terminology. They designated angiosperms as Magnoliophyta and the dicots and monocots as Magnoliatae and Liliatae respectively. These authors themselves are not very optimistic about the adoption of their new terms, for they realize the difficulty of replacing old terms which have gone into common usage for such a long time. Yet perhaps they were keen to point out the weakness of the old

terms. However, in our opinion, type-based terminology, especially for larger groups of plants, will suffer from the same limitations and consequently will expose itself to the same criticism as the character-based one. *Magnolia* and Lily, for instance, cannot well serve as types for the entire dicots and monocots respectively. So it is best to retain the old terminology. After all a term means what we want it to mean. Only we have to be constantly conscious of its limitations.

As to the division of angiosperms, if our pioneer taxonomists had access to all the numerous and varied species known to science today, and if they had available to them all the modern information gathered through collective efforts, perhaps they would have suggested a different classification and may have recognized more than two major classes. Therefore, the fact that most of our modern systematists have accepted the two major classes is not so much for the accuracy of the approach as for the convenience, which in fact is the primary concern of any classification. Any modification of this at this juncture, howsoever justified it may be, may create chaotic conditions without commensurate advantage.

Lotsy (1911) is believed to have proposed a system that differs from the previously proposed ones in assuming close agreement between the Araceae and the Piperaceae on the one hand and the Helobiae, etc., and the apocarpous Polycaricae on the other. Emberger (1960) carried this idea a little further by including the dicotyledonous Piperales with the Arales in the monocots. This is another concrete indication of the inherent weakness of our classification.

Monocotyledons, the group of lilies, orchids, cereals, bananas, palms, bamboos, etc., constitute a small class, comprising 5 sub-classes, 19 orders, 65 families and about 50,000 species (Cronquist, 1981). Perhaps one single factor that has been responsible for an all-around dwarfing of their status and also the various modifications in their habit and structure, is the absence of secondary growth in their axis. The usual angiosperm features that they share with the other class are:

1. The enclosure of the ovules within the ovary and pollination through stigma.
2. No free-nuclear stage in embryo development and reduction in size of male and female gametophytes.
3. Formation of mostly triploid endosperm through double fertilization.

4. Characteristic organization of shoot apex and its differentiation into tunica and corpus.
5. Occurrence of companion cell in phloem derived from the same mother cell as the sieve tubes.

Enclosure of the ovule within the ovary wall is perhaps the first known difference between angiosperms and gymnosperms. This ovule protection from the very beginning has, no doubt, provided the angiosperms with great evolutionary advantage. But in gymnosperm too ovules are exposed to outer atmosphere only upto the pollination time and thereafter they get well protected and enclosed by the up-growing adjacent scales. Then with regard to pollination, Endress (1980) has brought to light a very interesting situation where pollen, in some species of Monimiaceae, instead of being deposited on the stigma, germinates directly in the micropyle. further, the occurrence of 8-nucleate embryo sac and the double fertilization in endosperm development in a vast majority of cases in both the dicots and the monocots is often considered important in emphasizing affinities between the two sub-classes.

We agree that the very pains-taking discoveries and researches of many embryologists, especially those of Professor P. Maheshwari (1950) and his fellow workers, have been useful in solving many taxonomic and morphological problems, but these at the specific or generic levels only. At higher levels they may be as ineffective as those derived from several other areas of research. For instance, Onagrad type of embryo development occurs in such diverse families as Onagraceae, Magnoliaceae, Fagaceae, Rosaceae, Scrophulariaceae, etc. Similarly, Asterad type and other types may occur in very different and unrelated families. So, what is important is not the individual character but a syndrome of characters or a group of correlated characters (see Sporne, 1974).

### SOME PECULIARITIES OF MONOCOTYLEDONS

This group shows a number of features which are more or less peculiar to them. As the significance of some of them does not seem to be adequately appreciated, it may be worthwhile to focus some attention on them before commenting on the origin and evolution of the group.

**1. Habit and Habitat:-** Monocots are mostly herbaceous or shrubby. Essentially they are terrestrial but members of certain families, particularly those of Alismatidae, Juncaceae, Typhaceae, Cyperaceae, etc.,

have migrated to aquatic habitat, a fact most clearly betrayed by their flowers that, even in some submerged forms, stand out of water (see Sculthorpe, 1967). As members of very diverse and unrelated families, both in dicots and monocots, have taken to aquatic habitat, it is assumed that invasion into water took place several times in the evolutionary history of these plants. It has been estimated that of the major aquatic families, 18 belong to monocots and only 10 to dicots (Cook *et al.*, 1974).

Judging from the number of their species, monocotyledons exhibit a fairly rich variety in growth forms. This is revealed by features like sympodial and continuous modes of growth, restricted increase in thickness of stems, restricted branching, rapidly elongating slender internodes, occurrence of adventitious roots, leaf sheath, intercalary growth, production of resting organs like rhizomes, corms, stolons, bulbs, bulbils, etc. These under-ground and aerial organs survive when the rest of the plant dies down under adverse conditions. This seems to account for the infrequency of annual habit in this group. Outside the tropics, they are represented by small plants (Good, 1966).

The dendroid habit in monocots is met within certain groups of plants like screw-pines, agaves, bamboos and palms. The palms are, perhaps the most fascinating of all land plants. Corner (1966), who has chased them throughout the tropics of the world, describes a columnar stem crowned with giant leaves as an example of an ideal and perfect plant. *Calamus*, *Philodendron* and *Monstera* are amongst the giant climbers and *Zostera*, *Phyllospadix* and *Passiflora*, with their long internodes and leaves, are the only angiosperms which compete with marine algae (Stebbins, 1974).

In the largest of the monocot family, Orchidaceae, many of the species are epiphytic with aerial roots, whereas the terrestrial species have mycorrhizal roots. Members of the Triuridaceae are saprophytic with micotrophic non-green habit.

It will be seen that much of this variety in habit is attributable, directly or indirectly, to the complete or virtual absence of vascular cambium of the normal dicotyledonous type from all monocots (see also Holttum, 1955).

**2. Structure and Morphology of the Leaf:-** Leaf in monocots, like its counter-part in Pteridophytes, is generally a more important structural unit than it is in dicots. It is usually simple, sessile, with sheathing leaf-

base and parallel veins that converge at the tip. Although, these veins have lateral connections, they lack minute endings, so characteristic of the dicots. This distinctive venation pattern of the monocots is perhaps brought about by the peculiar mode of development in which both the apical and the lateral growth cease at a very early stage and is followed by intercalary growth at the leaf base. Most of the increase in leaf length is brought about by this latter growth. In grass leaves this intercalary meristem remains more or less permanently meristematic, enabling them to withstand frequent mowing or grazing. The leaves of grasses and the Cyperaceae are also unique in having stomata arranged in straight files in rows of one or two and all oriented in the same direction. In certain other monocots *e.g.* *Sansevieria*, *Allium*, *Juncus*, etc., the leaves may be cylindrical and unifacial.

Anatomically the monocot leaves are not so well differentiated into the palisade and spongy parenchyma. This is believed to be correlated with the chemical differences between the two classes, the monocots being usually sugar-leaved, and the dicots starch-leaved (Arber, 1925; p. 75).

The leaves in some palm species are often described as pinnately compound. But it must be pointed out that in their mode of development these are very different from similarly compound leaves of dicots. Such palm leaves, like any other monocot leaf, stop their apical growth at a very early stage. The intercalary meristem at the base becomes active and produces varying number of folds that mature distally. These folds split into leaflets from above downwards and form an apparently compound leaf which is merely superficially similar to its counter-part in dicots (*see* Corner, 1966; Periasamy, 1962 and others).

The morphology of monocot leaf has attracted considerable attention and still there exists considerable difference of opinion. Although, the leaf-skin theory proposed by Saunders (1922) could not get any support from any quarter, the phyllode theory, first proposed by A.P. de Candolle (1827) and stoutly supported and copiously documented by Arber (1925) in her monumental treatise—*Monocotyledons*, still appears to be popular with many authors in its original or somewhat modified form.

However, Arber (1950) had a second thought on this theory and appreciating the principle that 'a part can not be equal to the whole', she withdrew her support to it. With this withdrawal the theory should have stood as invalidated, for no other author, in my knowledge, has documented it to that extent, though several of them have commented on



its various aspects. This fact of Mrs. Arber's withdrawal of support to the theory was brought out in my review of this book (Puri, 1952). Again, it was made amply clear soon after her death, by her daughter, Muriel A. Arber in the Preface of the reprint edition of the 'Monocotyledons' brought out in 1961 by J. Cramer-Weinheim, Wheldon and Wesley Ltd., Hertz. But as far as I know, none of the modern authors, including Eames (1961), Cronquist (1968, 1988), Kaplan (1970, 1973)\*, Stebbins (1974), Sporne (1974)\*, Vassal & Maslin (1979), Sattler *et al.* (1988), etc. seem to have taken any note of this fact and they still refer to this theory as if nothing has happened to it.

Arber's revised interpretation of the monocot leaf is that it is "a fixation of the whole phylome at its pre-laminar stage". In other words, its development and differentiation ceased at an early stage resulting in a structure somewhat *analogous* to the very popularized phyllode of Australian acacias and yet being "potentially a whole phylome", it cannot be equated to the petiole which is "merely one element in this ultimate whole". This idea of fixation at an immature stage is not new. It is the same as neoteny and several authors, including Takhtajan have commented on it

Surprisingly enough, Arber did not elaborate this point any further and did not say anything about the basic issue, the phyllode met with in some Australian acacias, which perhaps provided the inspiration for the proposal of the phyllode theory for the monocot leaves. Perhaps this was the time when she lost all formal interest in botany and took to Indian philosophy, as she conveyed to me in a personal communication.

Any way, it is not difficult to envisage what she would have done if she had time to tackle the problem. She drew attention to a very sound principle, a part cannot be equal to the whole or that the whole cannot be equated to a part. It has wide implications and readily prompts a re-evaluation of the condition in the dicotyledonous acacias where phyllodes are generally believed to occur and from where analogy to the monocots seems to have been borrowed (Arber, 1950; p. 100). In these species, the first few juvenile leaves—three to five—are compound in as much as they comprise each a thin stalk bearing three, four or more compound pinna, the middle of the leaves usually having the largest number. Upward on the

---

\* Though Kaplan (1973) and Sporne (1974) cited Mrs. Arber's book in the bibliography, they do not seem to have noted her revision of her views on the subject.

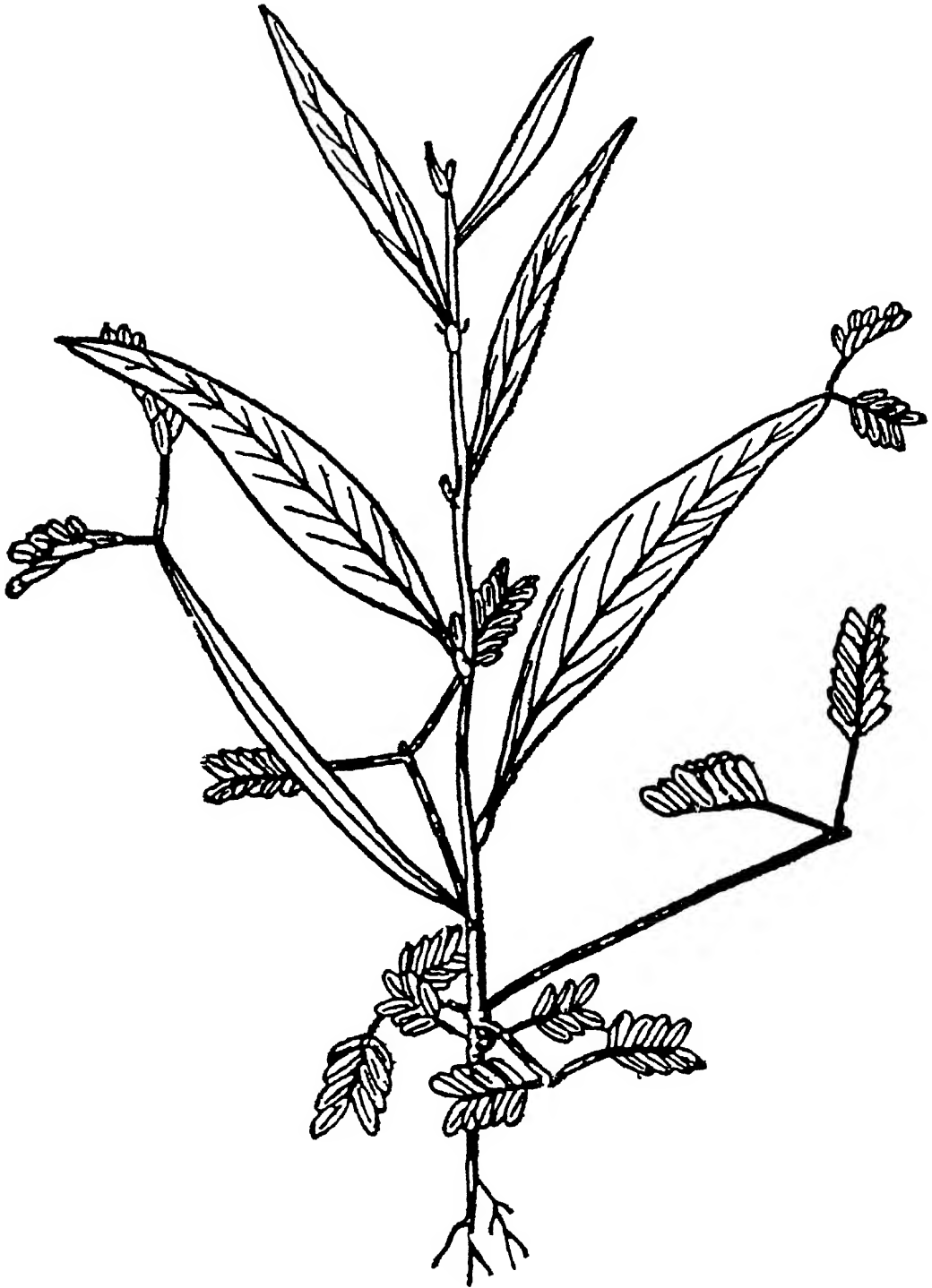


Fig 1 Juvenile pinnately compound leaves of Australian acacia. Adult leaves simple  
-- After Velenovsky (From Sinnott, 1961)

axis the number of pinna and pinnules decreases and the stalk becomes correspondingly longer and broader. Finally from the fifth leaf or so onward no pinna or pinnules develop and the adult leaves are represented by the flattened, elongated stalks that become isobilateral and end in a more or less pointed apex (Mucro), interpreted by Boke (1940) as "an abortive terminal leaflet." These are generally regarded as phyllodes, being modified petioles, an interpretation made all the more impressive by the appearance of young seedlings (Fig. 1).

The futility or absurdity of this interpretation becomes obvious if we apply the principle enunciated above to the situation in Australian acacias. It is well known that all leaves originate as small primordia on the axis. Depending upon the genetical and environmental factors, these primordia may develop into small or large, flattened or more or less cylindrical, simple or compound leaves. Some times they get modified into tendrils or thorns. But whatever they develop into they have to be regarded *morphologically whole leaves*, irrespective of their size, form or structure, and not just part of leaf. The question then arises as to how we should interpret these structures.

It is common knowledge that certain plants produce heterophilous leaves at different stages of development or under different environmental conditions. There may be two (dimorphic) or more (polymorphic) types of leaves in their life time. Many cases of dimorphic leaves are on record. In *Eucalyptus*, for instance, the first pair of leaves is horizontally oriented and dorsiventral, but all subsequent adult ones are more or less pendulous and bifacial. In species of *Hedera*, *Ficus*, etc., all vegetative leaves may be juvenile, which may continue for a long time. The adult leaves appear only in the reproductive phase and are different in shape and orientation (Fig. 2)

In aquatics or amphibious plants the phenomenon is more common. Arber (1919) reports that in *Sagittaria* the first leaves are thin and ribbon-shaped even when the plant is growing out of water and the upper ones are normal hastate type. Similarly, in several species of *Potamogeton* the floating leaves that rest on surface of water are relatively broad while submerged ones are ribbon-shaped. Some gymnosperms also illustrate the phenomenon.

It is, therefore, our contention that the condition in Australian acacias is best explained by assuming that the leaves here are of dimorphic type, the first few leaves or the juvenile leaves, being compound as they



FIG 2. Showing leaf dimorphism in *Hedera helix*. Flowering shoot with ovate, entire leaves, and a single leaf of the vegetative juvenile region ..... After Goebel (From Sinnott, 1961)

are in many other species of *Acacia* and the upper adult ones represented by flattened isobilateral stalk-like structures formed by excessive shortening of their developmental process. Such a situation is indeed tempting to reinvokethe doctrine of recapitulation. But this may be leading too much in a simple situation.

All this makes it clear that there is no necessity of envisaging a phylogenetic loss or suppression as Boke (1940) has done by suggesting that the so-called phyllode corresponds to the petiole and rachis of a pinnate leaf and that the "blade meristem or leaflet primordia which are a feature of ordinary foliage leaves have apparently been 'lost' in the phylogeny of the organ".

We, therefore, make hold to suggest that the expression phyllode is neither necessary for monocots, nor does it serve any useful purpose for dicots. In fact it is an incorrect and meaningless expression that be removed completely from botanical literature. While the expression phylloclade—leaf-like stem—is understandable, phyllode—petiole-like leaf is absurd and deserves to be discarded forthwith.

**3. Anatomy of the Axis:** Anatomy of monocots also presents several distinctive features, the most important being the "complete or virtual absence of a vascular cambium of the normal type from all monocotyledons" (Philipson 1971). Here the vascular bundles are scattered all over the ground tissue, never forming a vascular ring as in dicots. Also there is no inter-fascicular cambium. The effects of this absence of cambium have been far-reaching on the habit, structure and general physiology of the monocots.

Nevertheless, some monocot stems still show some limited increase in thickness, particularly in the arborescent forms. This may be effected usually in two ways: (1) Excessive cell divisions and elongation in the primary tissue, and (2) a different type of secondary growth peculiar to monocots only.

Columnar palm trunks, having almost uniform thickness all along, exhibit the first type of growth. There is no regular cambium involved here. Just beneath the growing point there are excessive cell divisions on the sides resulting in the formation of a shallow basin-like depression. The apical meristem proper is primarily concerned in the production of leaf primordia and contributes little to the stem tissue. The thickening of the axis is brought about by the activity of a meristem which is continuous beneath successive leaf bases and in which cell divisions are largely in tangential plane (Tomlinson, 1961). Secondary growth is considered to be the result of cell division and cell enlargement in parenchymatous ground tissue rather than due to any meristematic activity in a restricted region.

The second type of secondary growth occurs in some monocots with which all students of botany are well familiar. This occurs in plants like *Dracaena*, *Yucca*, *Agave*, *Dioscorea*, *Tamus*, etc. Here a secondary cambium develops out of the ordinary parenchymatous tissue and cuts off both parenchyma and vascular bundles. This may remain active through out the whole life of the plant, but the net thickness produced is not much. Since this occurs in several distinct and unrelated families, it is presumed

that this type of secondary growth originated repeatedly and independently in monocots.

Distribution of vessels in the wood is also of some interest as revealed by the studies of Cheadle (1953 and earlier work). Certain groups like Lemnaceae and some members of the Alismatidae do not have any vessels, the former lacks even tracheids. Other as reported by Cheadle may have vessels only in roots as in *Scheuchzeria*, Potamogetonaceae, certain Liliidae, Zingiberales, etc., and these have scalariform perforation plates. Alismataceae also have vessels only in the roots but surprisingly enough these have simple or scalariform end perforations. In palms, however, vessels occur in both root and stem and these have scalariform perforations. Cheadle considers them to be advanced in this respect.

It is significant to note that whereas vessels in dicots appear first (phyletically) in the secondary wood of the stem and then spread to other tissues and organs, they make their first appearance in the roots in monocots. This fact prompted Cheadle (1953) and others to suggest that vessels developed independently in monocots and dicots.

A detailed study of phloem in over 370 species of palms has revealed some interesting features. Parthasarathi (1968) has demonstrated that here the roots have compound sieve plates on very oblique to oblique walls, whereas the stems and petioles have compound to simple sieve plates on very oblique to transverse walls. Further, it is also noted that the sieve elements here have very long life, that may involve a century or even more, with little change in structure.

Branching in aerial stems of palms is rather rare but when it occurs it is of a distinctive nature. In *Hyphaene*, for instance, it is dichotomous and resembles that of *Pandanus*. Branches may arise leaf-opposed in rattan palms (Fisher and Dransfield, 1979) and non-axillary in two other species of palms (Fisher *et al.*, 1987). Tomlinson (1973) also reports leaf-opposed branching in some other monocots. All this is in marked contrast with the condition in dicots.

The structure and organization of stomatal apparatus also present significant differences. Whereas in dicots the two guard cells contain well developed chloroplasts and are surrounded by ordinary epidermal cells which in some cases become slightly modified, in many monocots these are smaller and contain chloroplasts that are poor in chlorophyll. Besides they are surrounded by two, as in Alismataceae, Juncaceae, Cyperaceae,

Gramineae, etc., or four or more as in Commelinaceae, Araceae, Palmae, etc., distinctive subsidiary cells. Liliales mostly do not have any subsidiary cells (Stebbins and Khush, 1961).

**4. Inflorescence and Flowers:** The type of inflorescence is also an important distinctive and unifying feature of monocots. A vast majority of them have inflorescence that are essentially racemose, spicate or some modifications thereof. Cymose type, which is common in primitive dicots, occurs only in some Alismatales and in palms, where they are enclosed in special structure, spadix (see Stebbins, 1974).

The size and composition of inflorescence in monocots varies considerably. The terminal inflorescence of *Corypha umbraculifera* may be about 10m high and 1 m thick at the base and may contain about 6,000,000 flowers; *Xanthorrhoea* (Liliaceae) is believed to have millions of flowers in a contracted 2m long panicle; while a single *Typha* inflorescence is estimated to have 300,000 flowers (see Eames, 1961).

The flowers of monocots are normally trimerous as against the tetramerous or pentamerous ones in dicots. However, as usual there are exceptions on both sides. They usually have smaller number of parts and may be small and inconspicuous as in the Cyperales, Palmae, etc., or large and showy as in the Liliales and Orchidales, etc. The perianth comprises usually six tepals in two whorls, which may or may not be differentiated. In *Potamogeton*, however, the flower has only four clawed tepals and four stamens opposite to them and four free carpels. In the past these flowers were sometimes interpreted as simple inflorescence comprising four staminate flowers surrounding one or more pistillate flowers in the centre. But this has been shown to be incorrect (Singh, 1965; Sattler, 1965). The Lemnaceae have the distinction of having the smallest plant body and the smallest flowers. Both the male and female flowers are without any perianth and consist of a single stamen and single, unilocular ovary (see Maheshwari and Kapil, 1963). In some other taxa the flowers have undergone much structural modifications as a result of adnation and cohesion. Some special structures as corona (Amaryllidaceae), Labellum (Zingiberaceae), column or gynostegium (Orchidaceae), etc., which were in the past erroneously interpreted are products of these processes (see Puri, 1952).

The gynoecium is apocarpous or syncarpous, sometimes with septal nectaries. Unlike the condition in dicots, the syncarpy here is mostly

associated with epigyny, which is believed to have originated in different ways and independently in several lines.

**5. Pollen and Pollination:-** The mode of formation of pollen grains is uniform throughout the angiosperms except in the Cyperaceae, where instead of the usual four, a pollen mother cell produces only one pollen grain. But a basic difference between monocots and dicots is seen in connection with the number and location of the apertures in the pollen grains. Whereas in monocots, as also in cycads, pteridosperms and Bennettitales, the pollen is uniaperturate (monocolpate or monosulcate) and the pore is located at the distal end, in dicots it is essentially triaperturate (tricolpate) and the furrows containing the apertures radiate from the proximal or inner end. Each of these two main types are believed to have produced multiaperturate and non-aperturate condition in their respective sub-classes and gave rise to other more complicated forms, the most evolved being the porate type.

In a majority of the angiosperms, pollen grains are shed at the two-celled stage, the tube cell and the generative cell. In other cases it is three-celled, the generative nucleus having divided before shedding. Such a situation among others, exists in the Alismatidae and is considered as advanced for it does represent further reduction of the gametophyte, so characteristic of all vascular plants. It is also believed to have strong survival value in plants that produce flowers under water.

Almost all modes of pollination are met with in monocots, but wind pollination or anemophily appears to be most common occurring in palms, pandanas, grasses, sedges, etc. In palms, however, Hendersen (1986) reports three syndromes—cantharophily (beetle), mellitophily (bee) and myophily (fly)—and considers anemophily as uncommon and derived. Gramineae and Cyperaceae, that are very highly evolved taxa, are considered to have been derived from entomophilous ancestors (Faggri and Pijl, 1966; Cook, 1988). True hydrophily is reported to be much more common in monocots. Les (1988) reports that out of 18 submerged angiosperm taxa having hydrophily, 17 are monocots. Highly specialized devices for pollination are met with in such taxa as *Vallisneria*, *Yucca* Orchidaceae, etc.

As to the entomophily, that is quite prevalent in Alismatales, Liliales, Orchidales, etc., Stebbins (1974) recognizes two conspicuous differences between the trends towards specialization for insect pollination in monocots and dicots. In the first place the differentiation of the perianth



into a protective calyx and a coloured attractive corolla is much less frequent in monocots than in dicots. He attributes this to weak stems which in turn are due to lack of vascular cambium in monocots. Another point refers to zygomorphy and epigyny which association too is much less frequent in monocots.

**6. Endosperm and Embryo:** Endosperm is the main source of nourishment for developing embryo and later in some cases for the germinating seed. Its formation is a result of the so-called double fertilization, an important peculiarity of angiosperms. Embryologists recognise three endosperm types, nuclear, cellular and helobial. Whereas nuclear type is peculiar to dicots, occurring only in the Araceae and Lemnaceae amongst the monocots, the cellular and helobial types are more frequent in the monocots. It is estimated that out of 17 families having helobial type, 14 are monocotyledonous (*see* Bhojwani and Bhatnagar, 1979).

The presence or absence of endosperm in mature seeds and the nature of its food reserves are important distinctive features. Alismatidae, for instance, are mostly non-endospermic and whenever the endosperm is present it is non-starchy, Commelinidae, on the other hand, have mostly starchy endosperm, whereas Arecidae have mostly fats, oils, proteins, hemicellulose, etc., stored in their endosperm.

The embryo in the Alismataceae, Butomaceae, etc., is simple, cylindrical and has little differentiation of organ, the sheathing cotyledon appearing terminal. In most other cases, excepting Orchidaceae, the embryo is well differentiated. Notwithstanding the primitive *Degeneria* having three or even four cotyledons and several other exceptions (*see* Eames, 1961), majority of the angiosperms have one or two cotyledons in their embryos and it is essentially on this feature that the class is divided into two sub-classes, monocots and dicots. The question which has been very hotly debated is the relationship of a single cotyledon of the monocots to two cotyledons in the dicots (Fig. 3).

In the 19th century monocots were generally considered as more primitive than the dicots and preceded the latter in any classificatory arrangement. As a consequence the single cotyledon was also considered as primitive and it was often suggested that by splitting, the single cotyledon gave rise to two cotyledons of the dicots. But now with more information at our disposal, tables have turned and monocots are generally considered more advanced and treated after dicots. Even Engler & Prantl's

system has now accepted this position (Melchior, 1964). Hence the theory of splitting is no longer in vogue and need not be discussed here.

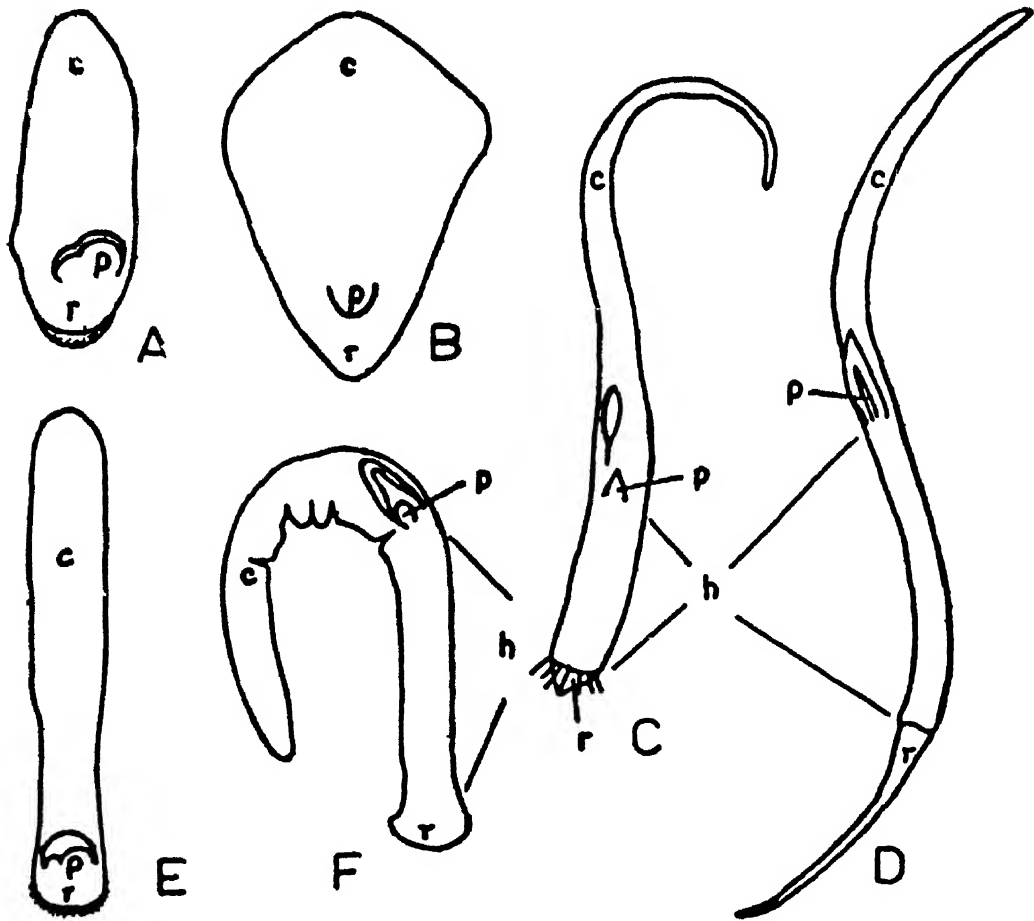


FIG 3 Various forms of simple monocot embryos. (A) *Arum*, (B) *Carex*, (C) *Butomus*, (D) *Alisma*, (E) *Sparganium*, (F) *Sagittaria*.— After Tschirch (From Eames, 1961)

This leaves two prevalent views in the field. According to one the single cotyledon of the monocots represents one of the two cotyledons of dicots. Among others, Eames (1961) is a strong exponent of this view and has discussed it in great details. The other view elaborated by Sargent (1903) and supported by Johansen (1945) and in modified forms by Cronquist (1968, 1988) and Stebbins (1974) is that the single cotyledon of the monocots is the product of congenital fusion or concrecence of

petioles of the two original cotyledons, the blades having been reduced or suppressed. The two views are almost equally balanced as far as the number and stature of the supporters are concerned and it is difficult to choose one from the other. But must one choose from them alone?

This entire controversy has been very ably discussed by several authors including Arber (1925), Eames (1961), Sporne (1974) and Stebbins (1974). It is, therefore, not necessary for us to cover this area here. Suffice it will to pose the question. Is it at all necessary to interpret the single cotyledon of monocots in terms of the two cotyledons of the dicots? We are one with Arber (1925) and Sporne (1974) believing that it is not at all necessary, just as it is not necessary in the case of gymnosperms. We do so notwithstanding Stebbins' rather harsh comments, dubbing all idealistic morphologists as those who "basically did not believe in either natural selection or evolutionary continuity" (Stebbins, 1974; p 328) and in another context as believers in "special creation". At this stage we would only wish that Professor Stebbins had better appreciation of idealistic or classical morphology\*, which, after all, is not antagonistic to evolutionary morphology (see Puri, 1952). In fact it is much older and its objective is quite different. While idealistic morphology deals primarily with the organism, Darwin's natural selection is only concerned with the relationship between the organism and its environment (see also Gutmann & Peters, 1973).

Cotyledons are just like any other leaves on the mature axis. In monocots every node bears only one leaf with broad sheathing base occupying a major portion of the nodal circumference and leaving little room for any other leaf to develop. The single cotyledon of the monocots is a homologue of this foliage leaf and nothing else (see Burger, 1981). Those who disagree overlook this point and seem to have become hypnotized, as it were, by their own terminology into thinking that cotyledons are some special organs *sui generis* and not the first leaves of plant (Arber, 1925; Sporne, 1974).

---

This reminds me of a prophesy that Mrs. Agnes Arber made while acknowledging the receipt of a few separates of my review of her book—*The Natural Philosophy of Plant Form*—(Puri, 1952). She wrote that white Indian scientists, with their usual background, can appreciate her writings, her Western colleagues would not

With a view to shed some light on the subject from another direction, Swamy and his associates studied the ontogeny of the single cotyledon in great details. In summarizing this work Swamy (1979) asserts that both the cotyledon and the epicotyl are organised in mutually adjacent positions in the derivatives of the terminal cell. One important variation, however, concerns the relative volume of the embryonic shoot apex that becomes involved in the organization of the cotyledon. In some taxa it is equal while in others a larger proportion of the shoot apex is involved in the development of the cotyledon. Swamy (1962) contends that early stages of development of the embryo in dicots and monocots are not identical.

### EVALUATION OF CHARACTERS

We have discussed here just a few selected characters. All of these obviously are not equally important. So far one feature, the structure of the embryo having one or two cotyledons has received the maximum importance and is made the main basis for dividing angiosperms into monocots and dicots. This was done at a time when our information about other features was very meagre and scanty and when taxonomists were in dire need of having some sort of classification as early as possible.

Now that other features also have been studied in detail and we have good deal of insight into them, it is desirable that we have a re-evaluation of all the important characters and determine, if possible, their relative value and significance. Of course, there are no definite guide lines for doing so. Every assessment will be more or less subjective. Besides, same character may have different values in different situations. Nevertheless, we are inclined to believe that by and large a character which is of very wide occurrence and which affects the general life of the entire group in a substantial manner should be considered as important.

The feature that meets both these criteria admirably well comes, in our estimation, from anatomy, the absence of vascular cambium from monocots. "this complete, or virtual, absence of a vascular cambium of the normal type from all monocotyledons and the restriction of any method of increasing thickness of stems and roots to very few species has had a profound effect on habit in this large group of plants." (Holttum, 1955; Philipson 1971; p. 112). Besides habit, it also restricts many monocots to particular habitats. It also affects the mechanism of pollination in some cases. In fact, the monocots are what they are mainly due to this particular

feature. No other feature, even though it may be equally prevalent, has influenced the make up of this group to the same extent. We are therefore, inclined to consider this feature as of greater importance than the possession of a single cotyledon. As a corollary it follows that in this particular respect gymnosperms are more akin to dicots than are the monocots. However, at this stage we have no intention of upsetting the *status quo*, nor do we consider ourselves to be competent to do so.

### ORIGIN OF MONOCOTS

Many contemporary evolutionists, including Cronquist, Takhtajan, Thorne and Stebbins, hold that the monocots have been derived from primitive dicots. However, Stebbins (1974) further adds that they have neither originated from any living order of dicots nor from any other group of non-angiospermous seed plants. Obviously he means that they are derived from some primitive, now extinct dicotyledons about which we know nothing directly.

Burger (1977), however, believes in close similarities between Piperales and Nymphaeales on the one hand and monocots on the other and proposes a hypothesis for deriving trimerous flowers from the piperalean ancestors. He envisages that simple flowers like those of the Chloranthaceae "came together by the loss of internodes to form the three-parted flower of the Piperales and many monocots".

Perhaps the strongest evidence in support of monophyletic origin of dicots and monocots and their close relationship comes from the occurrence of 8-nucleate embryo sac and the so-called double fertilization in the formation of the endosperm. These are no doubt important features. But there seems to be no difficulty in explaining them as products of parallel evolution in response to extreme reduction in the structure of the female gametophyte (*see also* Krassilov, 1977; p. 146). This becomes all the more plausible when we see that characters like vessels, closed venation pattern, inferior ovary, multiaperturate pollen, etc., are generally believed to have been acquired independently in various groups of angiosperms. Further, it is interesting to note in this connection that in *Hyacinthus* some pollen grains may develop further and organise into a structure similar to 8-nucleate embryo sac. And again, the two polar nuclei of the pollen grain may fuse together as in the real embryo sac, simulating double fertilization (*see* Bhojwani and Bhatnagar, 1979). This

shows that 8-nucleate condition and double fusion can be obtained elsewhere too under certain conditions.

The other point of view appears to emerge from the brief review of the more important morphological characters presented above. It reveals fairly clearly that in almost all cases-monocots stand apart from the dicots. Whether it is the habit, or general form, stem anatomy, root system, leaf development and morphology, stomatal structure and organization, type of inflorescence, floral organization, pollen structure or embryo structure, all present some striking differences which far out-weigh any recorded similarities between the two sub-classes.

Various approaches have been suggested here too. Khokhrjakov (1975, quoted from Krassilov, 1977), for instance, compiles evidence to suggest the derivation of monocots from the gymnospermous ancestors. Burger (1981), on the other hand, revives the old view that monocots are more ancient than the dicots and that they have originated from pteridospermous ancestors. He especially emphasizes the similarities between the embryo and the root systems of monocots and pteridophytes.

However, we are more impressed with the suggestion of Krassilov (1977) that the angiosperms as a whole, like any other major group of organisms could not have arisen "ready-made" as it were. The ancient forms acquired certain angiospermous characters to begin with and then gradually through the processes called angiospermization, attained all the important attribute of the group to become full-fledged angiosperms in the Upper Cretaceous period. Cronquist (1965), perhaps, also subscribes to a similar view when he emphasises that "there was no inherent point of time or morphological change at which we could say that *now and only now* the group has become angiospermous". Hughes (1977) also holds a similar opinion.

In this context it may be worthwhile to draw attention to the significant observations of Klaus (1979) and Cornet (1980)—both quoted from Muller (1981), reporting angiospermoid features in certain Triassic probably gymnospermous pollen. If these findings are confirmed they will lend further credibility to Krassilov's view point.

Krassilov further distinguishes three groups of Mesozoic seed plants, Czekanowskiales, Caytoniales and Dirhopalostachyaceae, as major sources of angiosperm characters. He believes that the monocots have

arisen from Czekanowskiales and the dicots from the remaining two groups.

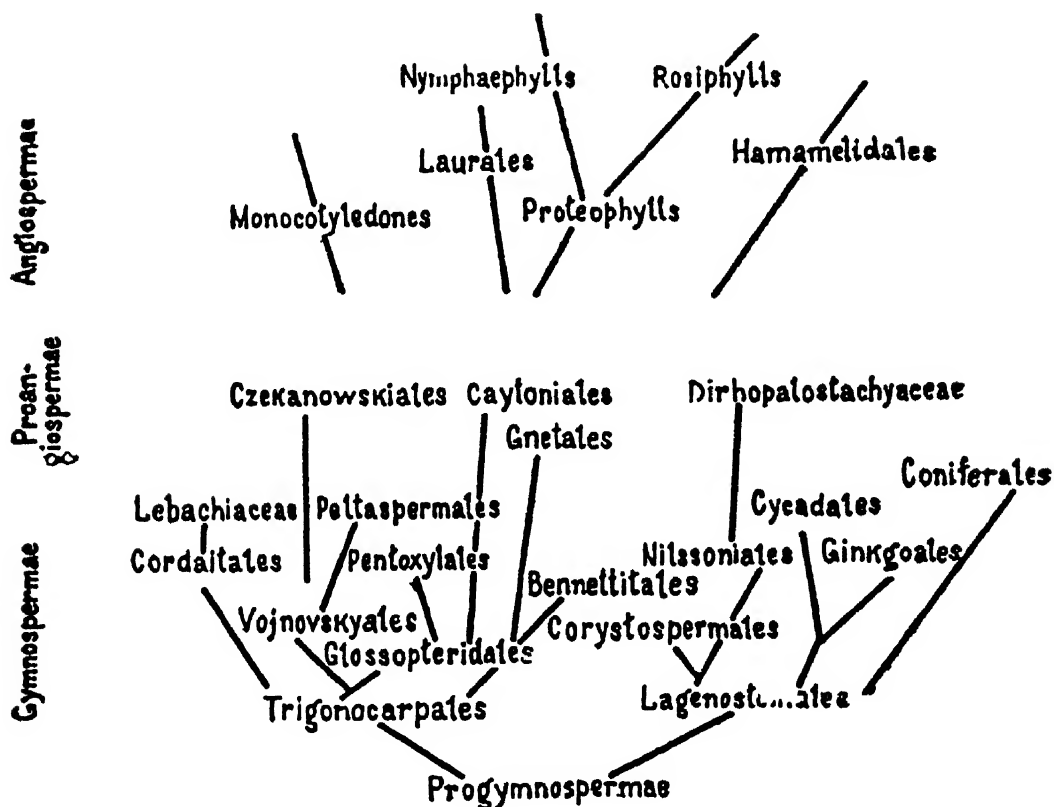


Fig 4 Putative relations of the proangiosperms to the gymnosperms and angiosperms

We do not mean to suggest that this hypothesis is satisfactory on all counts, even Krassilov does not claim that. But we do want to emphasize that this is the most plausible type of approach to the problem, it has elements of synthesising the various view points on the subject put forward during the past fifty years or so.

## INTER-RELATIONSHIPS AND EVOLUTION OF VARIOUS SUB-CLASSES

Several of recent taxonomists have recognized four sub-classes in the monocots, 1. Alismatidae, 2. Arecidae, 3. Commelinidae, and 4. Liliidae. More recently, however, Cronquist (1981, 1988) has separated Zingiberidae from the Commelinidae thus recognizing five sub-classes.

Every one of these sub-classes have some primitive and some advanced features and none can be regarded as basic or ancestral to another.

Alismatidae, for instance, are mostly apocarpous and are commonly held to be primitive, but they are advanced in having perianth differentiated into calyx and corolla, trinucleate pollen, that in some cases are multiaperturate with elaborate apertures and non-endospermic seeds. They are, therefore by no means primitive. Whatever similarities have been emphasized with the Nymphaeales in the past are all superficial and have no morphological basis (*see Tomlinson, 1982*).

Arecidae have four orders of which the two largest ones, Arecales (Palms) and Arales have about 3500 and 1800 species respectively. They mostly have arborescent habit and broad petiolate leaves. With the exception of Arales, all have vessels in all vegetative parts; endosperm is nuclear or cellular (Arales); floral reduction common; ecologically very variable.

Commelinidae have seven orders of which the Cyperals (12,000) and Brumaliales (1700) are the largest and Typhales (15), the smallest. Here the more primitive flowers are differentiated into calyx and corolla but more advanced ones may be reduced as in the Gramineae and the Cyperaceae and may be mainly wind pollinated. Pollen is trinucleate and endosperm nuclear. The group is very variable ecologically.

Zingiberidae have about 2100 species with Zingiberaceae and Musaceae having 1300 and 70 species respectively. They have pinnately-veined leaves, irregular flowers and inferior ovary and vessel mostly confined to roots.

Liliidae have two orders, Liliales (7700) and Orchidales (20,000), flowers with separate sepals and petals, petaloid, syncarpous, seeds either non-endospermic or endosperm with food reserve mostly other than starch; pollen mostly binucleate, well suited for insect pollination.

Recently Muller (1981) has given another exhaustive review of fossil pollen records of extant angiosperms wherein he gives a slightly different picture of the arrangement of various sub-classes of monocots on the basis of occurrence of pollen. Arecidae pollen is shown to be the oldest occurring in Maestrichtian (69 million years) and Alismatidae ones are shown to be the youngest (Upper Miocene, 11 million years). Liliidae and Commelinidae are shown to be of Upper Eocene (44 million years) and



Palaeocene (65 million years) periods respectively. This indicates that even the sequence of the various sub-classes is not well established so far.

Daghlian (1981) has given an excellent review of the geological history of the monocots. He distinguishes three phases in this evolution. In relation to palaeo temperature data. The first phase extends from Aptian-Albian to Maestrichtian and it is in this that the monocots make their first appearance in the form of isolated leaves and pollen. As such these can hardly be assigned to any modern forms. In the second phase that extends up to Eocene-Oligocene, the Cretaceous cooling having reached its maximum again warmed up and many recognizable forms such as palms, Zingiberaceae, Cannaceae, Musaceae, Cyperaceae, Gramineae, etc. appeared. The third successive phase began with more climatic changes, involving temperature, light intensity, air composition, etc., that could not adjust, disappeared completely and many new forms appeared amongst the old surviving ones. In the vegetation that survived all groups were not equally represented. Alismatidae and Arecidae had about half of their families represented, whereas Liliidae was represented by only one family. This is perhaps the reason why all major groups of the monocots appear so isolated and disconnected with one another.

The unique palms deserve a special mention here. Corner (1966) who has spent considerable time and energy in studying them in their native homes, asserts that whereas palms are generally included in monocot, it is more correct to say that "monocotyledons are palm derivatives." Fossil history appears to offer some support for such an inference (see also Moore & Uhl, 1982; Khakhryakov 1975). The earlier report that *Sanniguellia* leaf described from Jurassic of Colorado was similar to palm leaf was doubtful and rejected by several authors. But more recently Tidwall (1977) studied seven more specimens of *Snniguellia* and he re-asserts its palm affinities. If all this turns out to be correct, palm becomes so far the only angiosperm to have existed in the Jurassic. Muller (1981), however, does not go so far but he does list the fossil pollen of palms as the oldest amongst monocots. And Corner would not find any difficulty in deriving all the different monocots from the types of palms.

It is now well recognized, and our brief review also lends credence to this fact, that the problems of origin and evolution of monocots and that of interrelationships of their various groups cannot be resolved by a study of modern flora alone, howsoever perfect that may be. The problem has been complicated by large-scale destruction of early Cretaceous

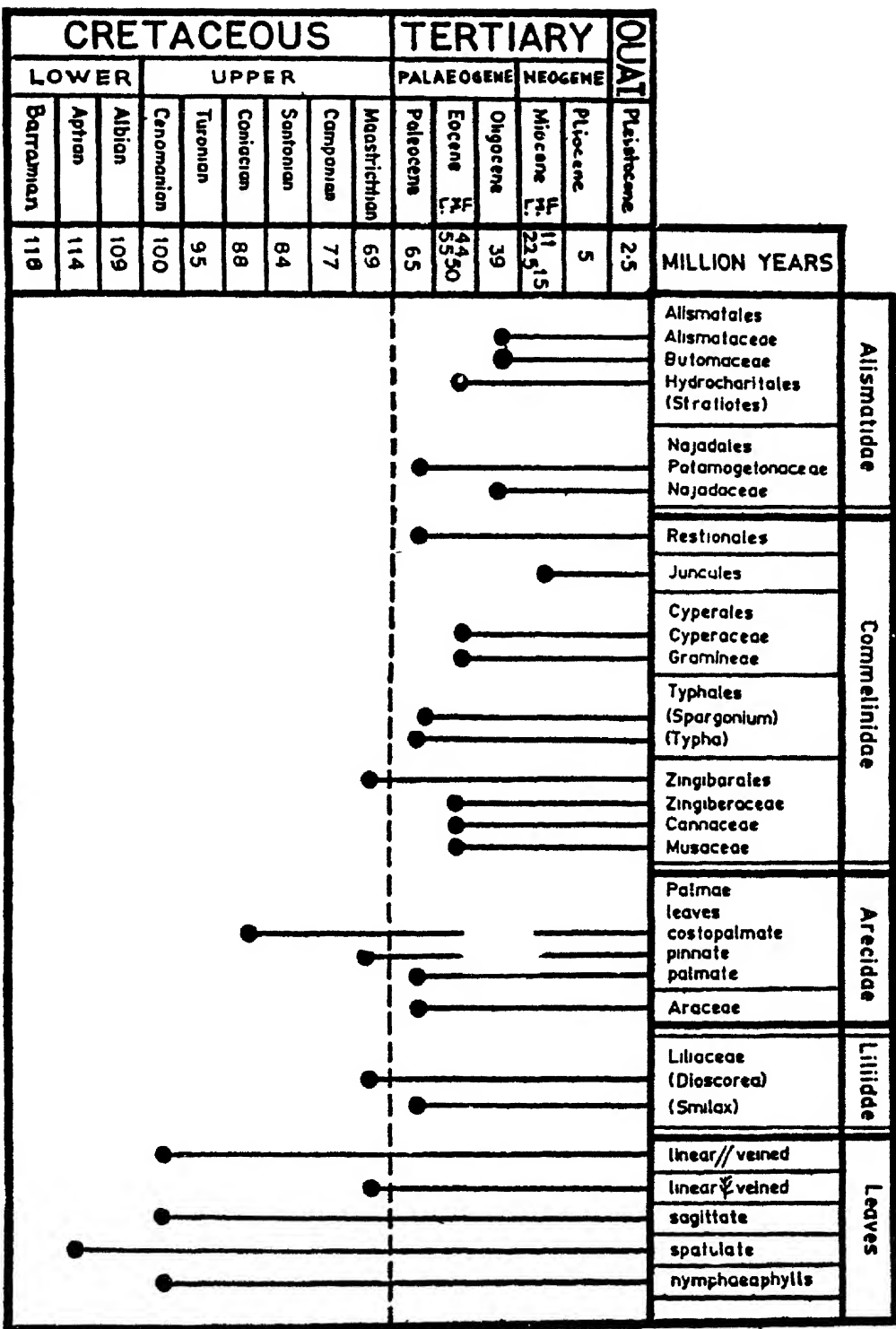


Fig 5 Stratigraphic ranges of monocot families —After Daghlman, 1981; Time scale after Muller, 1981

angiosperms and the remaining forms continuing as more or less unrelated taxa. Under such conditions it is only the fossil remains of the group through various geological periods that can help us. But macrofossils have believed all our expectations all these long years. Microfossils or fossils of spores and pollen, however, have already yielded substantial information and much of the insight that we have gained into the subject is through fossil pollen studies. More and more palaeobotanists are now turning to this aspect, helped as they are in a big way by the petroleum industry. We have, therefore, pitched high hopes again that sooner than later, 'missing links' between the earliest angiosperms and some undoubted taxa of some earlier group will be found and the "abominable mystery" that has baffled us so long, although not in the same virulent form as it did its author, can be resolved.

### ACKNOWLEDGEMENT

I am grateful to my colleague, Professor V. Singh, for going through the manuscript and making some useful suggestions, and also to Dr R Shyam, Head of the Botany Department, Meerut College, Meerut, for preparing the diagrams and the projection slides. I am also thankful to Dr D F Cutler and Dr Mary Gregory of the Jodrell Laboratory, Kew, for providing laboratory facilities during my stay in London in summer 1988 and also to Dr B S Venkatachala, Director, Birbal Sahni Institute of Palaeobotany, Lucknow, for some references and a geological time scale table. Thanks are also due to the Librarian, Kew Gardens Library, London, Dr R R Rao, Additional Director, B S I, Western Circle, Dehra Dun and the Librarian, F R I, Dehra Dun for permission to use their libraries.

### Selected References

1. Arber, A 1919, On heterophylly in water plants. *Am Nat* , 53 : 272-278.
2. Arber, A 1925 *Monocotyledons*. Camb, Univ. Press. Cambridge
3. Arber, A 1950 *Philosophy of the Plant Form* Camb. Univ. Press, Cambridge
4. Boke, N H 1940 Histogenesis and morphology of phyllode in certain species of *Acacia*. *Am J Bot.* 27 . 73-90
5. Burger, W C, 1977. The piperales and the monocots Alternative hypothesis for the origin of monocotyledonous flowers *Bot Rev.* 43 : 345-393.
6. Burger, W C, 1981. Heresy revived: the monocot theory of angiosperm evolution. *Evol Theory* 5 : 189-226.

7. Cheadle, V I 1953. Independent origin of vessels in the monocotyledons and dicotyledons. *Phytomorphology* 3 : 23-44.
8. Cook, Christopher D K, 1988. Wind pollination of aquatic angiosperms. *Ann. Mo. Bot. Gard* 75 : 768-777.
9. Cook, Christopher D K, B J Gut, E M Rix, J Schneller and M Seitz 1974. *Water Plants of the World* W. Junk, The Hague.
10. Corner, E J H 1966. *The Natural History of Palms*. Weidenfeld and Nicolson, London. 393 pp.
11. Cronquist A 1965. The status of the general system of classification of flowering plants. *Ann Mo. Bot. Grad* 52 : 281-303.
12. Cronquist, A 1968. *The Evolution and Classification of Flowering Plants*. Nelson, London.
13. Cronquist, A 1981. *An Integrated System of Classification of Flowering Plants* Columbia Univ. Press, New York. 1262 pp.
14. Cronquist, A 1988. *The Evolution and Classification of Flowering Plants* 2nd ed. The New York Bot. Garden, New York.
15. Daghljan, Charles P. 1981. A review of the fossil record of Monocotyledons. *Bot Rev.* 47 : 517-555.
16. Dahlgren, R 1981. Some conclusions from a comparative study of the monocotyledons and related dicotyledons orders. *Ber dt. bot Ges* 94 : 203-227
17. Doyle, James A. 1973. The monocotyledons. Their evolution and comparative biology. V. Fossil evidence on early evolution of the monocotyledons. *Quart Rev. Biol.* 48 : 399.
18. Eames, A J, 1961. *Morphology of the Angiosperms* Tata McGraw Hill, Faridabad.
19. Emberger, L 1960. In M. Chadeaud and L Emberger: *Traite de Botanique (Systematique)*. Tome 2, *Less Vegetaux Vascularies* Masson et Cie, Paris. 1539 pp.
20. Endress, P K 1980. Ontogeny, function and evolution of extreme floral construction in Monimiaceae. *Pl Syst Evol.* 134 : 79-120.
21. Faegri, K and PUL L Van Der 1966. *The Principles of Pollination Ecology*. Toronto.
22. Fisher. J B, C J Cott and A N Rao 1987. Non-axillary branching in two palms; *Eugessona* and *Oncosperma*. *Am J. Bot.* 74 : 613-614 (Abstr.).
23. Fisher, J B and J Dransfield 1979. Development of axillary and leaf opposed buds in rattan palm. *Ann. Bot.* 44 : 57-66.

24. Gutmann, Wolfgang Friedrich and D S Peters 1973. Konstruktion and Selektion: Arguments gegen einen morphologisch verkurzten selectionismus. *Acta Biotheo* 22 : 151-180
25. Henderson, A 1986. A review of pollination studies in the Palmae. *Bot Rev* 52 : 221-259.
26. Holttum, R E 1955. Growth habits of Monocotyledons—Variation on a theme. *Phytomorphology* 5 : 399-413.
27. Hughes, N F 1977. Palaeo-succession of earliest angiosperm evolution. *Bot Rev* 43 : 105-127
28. Johansen, D A 1945 A critical survey of the present status of plant embryology *Bot Rev* 11 : 87-107.
29. Kaplan D R 1970. Comparative foliar histogenesis in *Acorus calamus* and its bearing on the phyllode theory of monocotyledonous leaves *Am J Bot* 57 331-361
30. Kaplan, D R 1973. The monocotyledons: Their evolution and comparative biology. VII The problem of leaf morphology and evolution in the monocotyledons *Quart Rev Biol* 48 437-457
31. Kaplan, D R 1975 Comparative developmental evaluation of the morphology of unifacial leaves in the monocotyledons. *Bet Jahrb Syst* 95 1-105
32. Kokhrjakov A P 1975 *Somatic Evolution of the Monocotyledons* Nauka, Moscow 196 pp (In Russian)
33. Krassilov V A 1973 Mesozoic plants and the problem of angiosperm ancestry *Lethaia* 6 : 143-176
34. Krassilov, V A 1977. The origin of angiosperms. *Bot Rev.* 43 : 143-176.
35. Les D S 1988. Breeding systems, population structure and evolution in hydrophilous angiosperms. *Ann. Mo. Bot Gard* 75 : 819-835
36. Maheshwari P 1950. *The Embryology of Angiosperms* McGraw-Hill, New York
37. Maheshwari, S and R N Kapil 1963. Morphological and embryological studies on the Lemnaceae 1. The floral structure and gametophytes of *Lemna paucicostata* *Am J Bot.* 50 677-686.
38. Melchior, Hans 1964. *Engler's Syllabus der Pflanzenfamilien*. Ed 12, Vol. 2. Gebruder Borntraeger, Berlin.
39. Moorie, H E and N W Uhl 1982. Major trends of evolution in palms. *Bot Rev.* 48 : 1-69.
40. Muller, Jan 1981. Fossil pollen records of extant angiosperms. *Bot. Rev.* 47 : 1-142.

41. Palrwal, G S 1969. Stomatal ontogeny and phylogeny 1. Monocotyledons. *Acta Bot Neerl* 18 : 654-668
42. Parthasarthy, M V 1968. Observations on metaphloem in the vegetative parts of palms. *Am J bot* 55 : 1140-1168
43. Periasamy, K 1962. Morphological and Ontogenetic studies in palms. *Phytomorphology* 12 : 54-64
44. Philipson, W R 1971 *The Vascular Cambium* Chapman and Hall Ltd., London. 182 pp.
45. Puri, V 1951 The role of floral anatomy in the solution of morphological problems. *Bot Rev.* 17 . 471-553.
46. Puri, V 1952. Review of Agnes Arber 1950. "The Natural Philosophy of Plant Form" *Phytomorphology* 2 : 186-188
47. Sargent, E 1903. A theory of the origin of Monocotyledons, founded on the structures of their seedlings. *Ann Bot* 17 · 1-92.
48. Satler, R 1965. Perianth development of *Potamogeton richardsoni*. *Am J. Bot.* 52 : 35-41.
48. Scuthorpe, G D 1967. *The Biology of Aquatic Vascular Plants* Edward Arnold, London
49. Singh V 1965. Morphological and anatomical studies in Helobiae. II Vascular anatomy of the flower of Potamogetonaceae. *Bot Gaz* 126 : 137-144.
50. Sinnott. E W 1960 *Plant Morphogenesis* McGraw Hill Co., New York.
51. Sporne, K R 1974. *The Morphology of Angiosperms*. Hutchinson Univ. Library, London.
52. Stebbins, G L 1974. *Flowering Plants Evolution Above the Species Level* Edward Arnold, London.
53. Stebbins, G L and G S Kush 1961. Variation in the organization of the stomatal complex in the leaf epidermis of monocotyledons and its bearing on their phylogeny. *Am J Bot* 48 51-59.
54. Swamy, B G L, 1962. the embryo of monocotyledons: A working hypothesis from a new approach. In *Plant Embryology – A symposium*. pp. 113-123. CSIR, New Delhi.
55. Swamy, B G L 1979. Embryogenesis in *Cheristylis flabellata*. *Phytomorphology* 29 : 199-203.
56. Takhtajan, A 1969. *Flowering Plants : Origin and Dispersal*. Oliver and Boyd, Edinburgh.
57. Teslenko Y V 1967, some aspects of evolution of terrestrial plants. *Geologia i. geofizsa* (Novosibirsk) 11 : 58-64.

58. Tomlinson. P B 1961. *Anatomy of the Monocotyledons II. Palmae* The Clarendon Press, Oxford.
59. Tomlinson. P B, 1982. *Anatomy of the Monocotyledons VII. Helobiae (Alismatidae)*. The Clarendon Press, Oxford.
60. Vassal. J and B R. Maslin 1979. Phyllodes et phyllodisation dans le genre *Acacia*. *Bull Soc. Nat. tonlonse* **115** : 393-401.
61. Wu, Qi-Gen and D F Cutler 1985. Taxonomic, evolutionary and ecological implications of leaf anatomy of rhizomatous *Iris* species. *Bot J Linn Soc London* **90** · 253-303



Dharani Dhar Awasthi (b. 28 September 1922) did Ph.D.(1961) from University of Lucknow and Ph D (1962) from University of Colorado, Boulder, USA. He was INSA Senior Scientist (1987-90); Emeritus Scientist (CSIR) (1984-87), Botany Department, University of Lucknow; Reader in Botany, University of Lucknow.

Awasthi monographed a majority of macro-lichen and several micro-lichen genera of India. Monographed genus *Dirinaria* on world-wide basis. Described three new genera, over 70 new species, numerous subspecific taxa and new combinations.

Consolidated information in two publications on all the known macro- and micro-lichen taxa of India in the form of keys for their identification.

Awasthi was President of Indian Mycological Society (1982), and is Life Member, American Bryological and Lichenological Society and Society of Sigma Xi, USA; Fellow, Indian Academy of Sciences; Member, Indian Mycological Society. He is the recipient of Professor Panchanan Maheshwari Memorial Lecture Award (INSA) (1991), Honorary Member, British Lichen Society (1992); ACHARIUS Medal of International Association for Lichenology (1992).

*Dharani Dhar Awasthi was elected to the fellowship of the Academy in 1984*

**Address:** 38, Khursheed Bagh, Lucknow 226 004; Tel: (0522) 236686.



## THE LICHENS AND THEIR PECULIARITIES

DHARANI DHAR AWASTHI

Mr Chairman, Distinguished Scientists, Ladies and Gentlemen:

In this august assembly of specialists of different scientific disciplines, my talk pertains to a small group of plants— the lichens, and their peculiarities. For the benefit of the non-botanists it may be desirable to give a short historical sketch of the lichens with elementary information for which I apologise to the learned botanists here.

The word 'lichen' is of Greek origin. It was first used by Theophrastus in the third century B C for the superficial growth occurring on the olive trees. It is likely that it may then have encompassed the bryophytes as well. The term 'lichen' was reserved by Tournefort in 1700 A D for the group of plants as we know them now by calling it a genus *Lichen*. The foremost account of 80 species under the genus *Lichen* was published in 1753 in *Species Plantarum* by Linnaeus, who placed this genus in the 24th Class Cryptogamia Algae. The exact morphological nature of the lichens remained obscure for another over a hundred years, though a large number of new species and genera were described and classified in several publications during that interval by Acharius, Fée, Fries, Koerber, Schaerer, Massalongo, Tuckermann and Nylander. New terms for the specialised structures present in the lichens were also coined. The greenish cells seen in the lichen thallus were erroneously considered reproductive bodies and called 'gonidia'. This term survived almost till the mid-twentieth century, being used in the sense of the algal component of the lichen thallus; and presently called photobiont.

The exact morphological nature of the lichens was discovered in 1867 by the Swiss botanist Schwendener, who proposed the dual hypothesis for the lichen thallus. According to him the lichen thallus was constituted by a fungus and an alga, the former acting as a parasite on the latter. However, the biology of lichen thallus wherein a healthy growth, multiplication and reproduction took place like a single plant without the annihilation of the alga could not be explained by referring the fungus as a parasite. A few terms were proposed to explain the association of the two components but the term 'symbiosis' (also called conjoint life) proposed by De Bary in 1879 has received acceptance in general even now. In the

symbiotic association of the two constituents in the lichen, De Bary considered both the components to derive benefit from each other, though not necessarily to an equal degree.

Presently three types of symbioses are recognized. In the mutualistic (biotrophic) symbiosis there is a stable relationship and both the partners benefit. In the commensalistic symbiosis one partner benefits and the other is neither adversely affected nor benefited in any way. In the antagonistic (parasitic or necrotrophic) symbiosis one partner benefits at the expense of the other, which is often killed in the process. The latter two types of symbioses are ruled out for explaining the relationship between the two components of the lichens as they live harmoniously for a long period of time effecting growth, reproduction and performing normal physiological processes like the autotrophic plants. Lichens therefore can be considered to exhibit mutualistic symbiosis. But there are certain other organisms which are associated to provide mutual benefit e.g. the fungal association in the roots of certain conifer trees (called mycorrhiza), the algal association within the roots of *Cycas*, none of which are called lichens. It can then be envisaged that when an alga and a fungus are associated in mutualistic symbiosis it is to be called a lichen. But is it so? It has been found that certain fungi and certain algae live in an apparent symbiotic relationship but are not referred to as lichens. For example, the fungus *Mycosphaerella ascophylli* grows inside the tissues of the brown algae *Ascophyllum* and *Pelvetia*, and another species *M. apophlaeae* grows inside the tissues of the red alga *Apophlaea*. The fungal hyphae ramify throughout the algal hosts and develop pycnidia and ascocarps on the surface of the host. The algae also remain healthy and reproduce sexually. Although the physiology of these associations is not well understood, it has been observed that the above uninfected algae do not survive in nature. Similar situations have been reported in the associations of the fungus *Kohlmeyera complicatula* with the alga *Prasiola*, hyphomycetes fungus *Blodgettia confervoides* with the alga *Cladophora* and the fungus *Phaeospora lemaneae* with the alga *Lemanea*. None of the above are called lichens. The reasons for not calling them lichens are the endophytic (inhabitant) nature of the fungal hyphae, the structure of the association is that of the alga, and that the sexual reproduction of the alga is a persistent feature in the partnership. In general, the fungus of lichen is not endophytic, the structure of the lichen thallus is unlike the two partners, and the alga does not reproduce sexually. There are, however, exceptions in one or the other criterion (see later).

As the true biology of the lichen thallus has gradually been unfolded, it will be futile to give here the definition of the lichen by the earlier lichenologists and mycologists. In 1979 Alexopoulos and Mims defined a lichen as "an association of a fungus and an alga in which the two organisms are so intertwined as to form a single thallus". A couple of years later, Ahmadjian (1982a, b) defined the lichen as "an association of a fungus and the photosynthetic symbiont resulting in a stable thallus of specific structure", which, "does not resemble either symbiont in the free living (unlichenized) state".

However, in the well established lichen taxa *Coenogonium* (photobiont usually a *Trentepohlia*), *Racodium* (photobiont a *Cladophora*) and *Ephebe* (photobiont a *Stigonema*) the basic structure of the lichen thallus is that of the photobionts *Trentepohlia*, *Cladophora* and *Stigonema* respectively. There are few other examples in this category. The filaments of the photobiont are ensheathed by the mycobiont from outside to form the lichen thallus. Therefore, to improve the latter part of Ahmadjian's definition, Hawksworth (1988) has proposed "a lichen is a stable self-supporting association of a mycobiont and a photobiont in which the mycobiont is the exhabitant". Apparently Hawksworth did not take the haustorial development into consideration, as then this definition also becomes partly untenable in the light of the mycobiont not being strictly exhabitant when it produces haustoria into the photobiont cells in some lichen taxa. Hawksworth (l.c.) has also cautioned "Biologists in general and mycologists in particular need to recognize that a single definition which can unequivocally categorize each mutualistic fungus-alga relationship discovered is not achievable; where possible broader symbiological terminology should be adopted".

In contrast to the acceptance of the symbiotic relationship in lichens, Ahmadjian and Jacobs (1983) are of the opinion that the mycobiont is a biotrophic parasite maintaining a balanced relationship between the fungus and the alga and thus their association be called controlled parasitism. In the culture experiments it was seen that the mycobiont of *Cladonia cristatella* behaved as parasite on most of the algae. Few algal genera and species other than its own photobiont may form small lobules of a lichenized thallus, which also get parasitised later. The lichenization is determined by the physiological behaviour of the alga.

### *Morphological and Distributional Diversity of Lichens*

Thus defined, the lichens are remarkable perennial plants, which morphologically look and physiologically behave as a single biological unit. In general, the morphology of the thallus is different from either of the symbionts when growing in natural (unlichenized) state. The delicately balanced physiological processes in the symbiotic relationship can endure an immensely slow growth, probably the slowest in the plant kingdom. Lichens can grow on anything that is capable of remaining stable to support and sustain the extremely slow growth. Climatically lichens are distributed in the diverse geographical regions from hot tropics to frigid high mountains and arctics, from extreme xerophytic areas of the deserts to the very moist conditions, to the extent of growing on rocky substrates that are periodically or persistently submerged in water and from sea level to high mountains. The substrate is also much variable. Lichens are found growing rather luxuriantly on rocks, boulders, soil, decaying plant material, bark of perennial plants, on the surface of perennial leaves, on glass panes, on metallic structures, on insects (weevils in New Guinea), on shells of turtles etc. Presently the fungal component is called mycobiont and the algal or cyanobacterial (previously called Cyanophyceae algae) component photobiont. The mycobiont in the great majority of lichens is a member of the fungi Ascomycetes. In few taxa of lichens the mycobiont belongs to Basidiomycetes and Hyphomycetes. The estimates of the total number of lichen species from the world vary from 14000 to 20000. Diversity in the morphological nature of the thallus in so many taxa is naturally expected. The thallus exhibits a great deal of structural variation in its crustose, squamulose, foliose and fruticose forms. There is also a variation in the thallus being superficial or partially to completely developed within the substrate (bark or rocks), in which only the fructifications (ascocarps) may partially or entirely be superficial.

The present tendency of the lichenologists and the mycologists is to incorporate the lichens, also called lichenized fungi, within the fungal system of classification. Since the mycobionts of the majority of the lichens belong to the Ascomycetes, a perusal of that group indicates that 5 orders of them exclusively constitute the lichens, while in the remainder there are some which contain both lichen forming and non-lichen forming taxa. It can therefore be said that during the process of evolution of lichenization there have been attempts made by the different ascomycetous fungi to develop lichenized associations with the green algae or cyanobacteria and that a fairly good percentage of them were

successful in those attempts. That this must have been the case is also reflected on a perusal of the whole group of lichens, in which we find a continuum of associations of the two bionts from casual lichenization, loose association to the formation of the complex lichenized structures. Greater the complexity in the thallus, greater the domination by the mycobiont. Special modes of sexual reproduction and vegetative propagation have been contrived with the sole aim of survival in the lichenized capacity as survival of the mycobiont in the natural state was no longer possible.

Loose symbioses with algae or cyanobacteria not leading to lichenization or specialized morphological structures are reported in the diverse groups of fungi. In the Arthopyreniaceae the members of which generally inhabit the bark, species *Mycomicrothelia melanospora* is nonlichen forming. Species like *M. atlantica* has small patches of the photobiont *Trentepohlia*, while other species like *M. thelena* are consistently lichenized (Hawksworth 1988). Some species of lichen genus *Epigloea* lacking differentiated thallus have been called 'half - lichens'. About 45 taxa of conidial lichen forming fungi have been reported in which there is a stable association with the photobionts. Some of them, e.g. *Coniosporium aeroalgicola*, have been reported as 'semi-lichens'. In the lichen genus *Collema* (photobiont a *Nostoc*), colonies of *Nostoc* on lichenization by the mycobionts get characteristically modified in the formation of the different types of thallus. The cells of the two bionts are irregularly dispersed throughout the thallus which remains undifferentiated. In the advanced forms of lichens the major part of the thallus is made up of the mycobiont in various modifications, while the photobiont is stratified below a cortex formed by the mycobiont. It has been suggested that the differences in the morphology, anatomy, physiology, reproduction and to a greater extent the chemistry of the different taxa of lichens are the results of the interactions between the two bionts. The photobiont is supposed to be of greater importance as it is the photobiont that stimulates the mycobiont for the phenotypic expression of the thallus. The orientation and morphology of the lichen thallus depends on the photobiont so that the photobiont itself remains exposed to the sunlight as well as be protected by the cortical tissue. It was found that if a thallus was turned upside down, the photobiont layer disintegrates and a new layer is formed by the mycobiont on the upperside under stimulation of the errant photobiont cells present in the medullary region.

### *Number of Bionts in a Lichen Taxon*

Usually a lichen is formed by two bionts, one of which is a mycobiont and the other a photobiont. But in certain taxa of lichens in which the main or primary photobiont is a green alga, a cyanobacterium (*Nostoc*, *Gloeocapsa*, *Scytonema*, *Stigonema*) is associated as a secondary photobiont. In the genera *Placopsis*, *Pilophorus*, *Stereocaulon* and few species of *Peltigera* (e.g. *P. aphthosa*) the main photobiont is a green alga, but on the surface of the thallus there are present morphologically distinct structures which contain a cyanobacterium and are called external cephalodia. These cephalodia represent localized growth on the main thallus by the incorporation of the cyanobacteria by the hyphae of the mycobiont and do not have independent existence. In species of the foliose lichen genera *Solorina*, *Nephroma*, *Lobaria*, *Sticta*, *Pseudocyphellaria* which have a green alga as the primary photobiont, there are often found separate, delimited areas in the lower part of the medulla containing a cyanobacterium and these areas are called internal cephalodia. The reverse type of association or combination i.e., primary photobiont a cyanobacterium, secondary photobiont a green alga has not been reported. Functionally, cephalodia supplement the nitrogenous requirements of the lichen in which they occur, and can be compared with the role of the root nodules in the leguminous plants. The origin and anatomy of the cephalodial development had long been worked out in detail. Dughi had noted a relationship of the free living and attached form of cephalodia and pointed out that the mycobiont of *Dendroscopaulon* and species of *Sticta* and *Lobaria* could be the same. The morphological significance and the taxonomic implications of the cephalodia have been reviewed by James and Henssen (1976) and added their own observations to the problem. Intensive field studies revealed the occurrence of pairs of conspecific morphotypes, each of independent existence in certain taxa of lichens, in which the association of a cyanobacterium or a green alga with the same mycobiont could form morphologically different thalli. For example, the caulescent morphotype lichen *Dendroscopaulon* (with cyanobacterium photobiont) could bear at its extremities small lobules of the lichen *Sticta filix* (with a green algal photobiont). Conversely *Lobaria amplissima* (with green algal photobiont) could develop coralloid growth of the morphotype *Dendroscopaulon*. It was also found that the taxa *Sticta dufourii* and *S. canariensis* are the two morphotypes of the same mycobiont with a cyanobacterium or a green alga respectively. It was also observed that the two morphotype thalli of *Dendroscopaulon* sp. and *Sticta filix* were present

at the same place at different levels on two rocks and that apparently the environmental conditions favoured the development of one or the other morphotype. Morphotype *Dendriscocaulon* occurred in the shady niche with high humidity and *Sticta filix* in exposed sunny area with less humidity and composite thalli in the middle. It is also of interest to note that in all such cases, the morphotype *Dendriscocaulon* (photobiont a cyanobacterium) is always sterile and reproduces by vegetative propagules only. The green algal morphotype *Sticta filix* has a well developed and differentiated foliose thallus with ascocarps. On this basis, it has been suggested that the *Dendriscocaulon* type morphotype is advanced over the *Sticta filix* morphotype; the sexual reproductive system has been dispensed with in favour of the vegetative propagation, in which the bionts, particularly the mycobiont is better suited for survival. Were it so, the majority of the lichens would have now evolved to possess cyanobacteria as the photobiont. On the contrary majority of the lichens have a green alga as the photobiont, and have variously modified asexual or vegetative propagules. Awasthi (1975) had indicated that in the early stages of lichenization it were the cyanobacteria that may have been involved due to their primitiveness, and there must have been a shift from cyanobacteria to green algae as the latter are more efficient for photosynthesis. The presence of the cephalodia in certain lichens, is therefore a recapitulation of the ancestral habit.

It was observed that if the photobionts belong to two closely related species the thallus formed does not show morphological variation. It was discovered that in the lichen taxon *Xanthoria parietina* the photobiont could be either *Trebouxia albulescence* or *Tr. decolorans*. Similar associations are likely to be present in other taxa.

### *Nomenclature of Lichens*

It is an irony that inspite of 80 species having been described under the genus *Lichen* in Species Plantarum in 1753, the starting date of the botanical nomenclature, the genus *Lichen* is not valid to-day. The 80 heterogenous species had been transferred to other genera and the genus *Lichen* remained untypified. In addition, the International Code of Botanical Nomenclature stipulates that the name of a lichen taxon pertains to the mycobiont, the photobiont has to be named separately. For example, in the lichen taxon *Parmelia nepalensis*, the mycobiont is *Parmelia nepalensis*, while the photobiont is a *Trebouxia*. This is in utter variation to the plant names in other groups. There are thus as many additional taxa

of fungi as the number of lichen taxa. But the same does not hold good for the photobionts. Only about 30 genera of the photobionts (algae and cyanobacteria) are known. Production of several thousand lichen taxa by the species of the 30 genera of photobionts is easily conceivable if we remember that the same or the same type of photobiont may stimulate different mycobionts differently resulting in the variety of structural differences in the thallus and fructifications.

### *Physiology of Lichens*

The general physiological processes going on in the lichen thallus are somewhat different from an autotrophic plant. The mycobiont is a fungus and thus basically heterotrophic and subsisting on the food prepared by other organisms. The photobiont of the lichen synthesises the food during photosynthesis from the elements and simple compounds that are easily available from the environment. The mycobiont acts as a moisture holding structure for the maintenance of the normally aquatic photobiont and protects it from the excessive sunlight and desiccation. The synthesized food is partially used by the photobiont and partially gets transported by diffusion to the mycobiont in the form of sugars and polyols. The hyphae of the mycobiont are in close proximity of the photobiont cells, which are either clasped by the development of appressoria, or in some cases haustoria are also produced. The haustoria may either just break open the cell wall of the photobiont without penetrating the plasma membrane or rarely may also be intramembranous. But in spite of the presence of the haustorium within the photobiont cell, the latter is capable of cell division. For this interesting aspect, it has even been suggested that the presence of the haustorium induces the photobiont cell to divide.

In the obligately foliicolous lichens *Strigula* and *Raciborskiella* the thallus develops in the subcuticular region of the leaf, and thus is separated from the epidermal and underlying mesophyll or palisade cells by their walls. But interestingly the hyphae of the mycobiont do not involve or penetrate the epidermal cells of the leaf in search of food. The mycobiont is content with the frugal food synthesized by its own photobiont.

Lichens are adapted to survive in extremes of xeric conditions. They revive quickly with the availability of moisture from the humid atmosphere or dew during night. The photosynthetic and other metabolic activities are completed within a short period of time in the morning, and then there is a cessation of the physiological activities, the lichen thallus



becoming latent, but not die out. Lichens are also active photosynthetically at very low temperatures, even below freezing point in the cold arctic regions.

In comparison to the whole thallus, the amount of photobiont in the lichen thallus is generally very little. The food synthesized is consequently small, so that a rapid or vigorous growth is not possible. The slow growth of the lichens is apparently also due to the low rate of CO<sub>2</sub> assimilation, non-availability of proper nutrients, water and environmental conditions. All these allow a short period of optimum metabolic activities which lead to slow growth. It is estimated that in the majority of lichens the growth varies from 0.1 to 10 mm per annum. This extremely small yearly growth is probably also necessary for the continuation of their delicate symbiotic relationship. The growth takes place peripherally (crustose and foliose forms) or apically (fruticose forms). There is none to very little thickening of the thallus in the central or basal part than what developed initially. In addition, irrespective of the region or age of the thallus or its part, the cells of the two bionts remain alive and physiologically active for a long period of time. Lichens persist for several years, and much more than any of the other associations. The foliose and fruticose lichens are known to be living for 15-80 years, while the thalli of crustose lichens like *Rhizocarpon* are estimated to be living for over a hundred to few hundred years. The hymenium of the ascocarps remains active in the production of asci and ascospores for several years in contrast to the hymenium of the non-lichenized fungi. The attainment of maturity after a period of time does not necessarily mean the beginning of senescence in lichens. The foliicolous lichens are, however, exceptions as they usually complete their life cycle within one or few years. This is also a biological adaptation as even the perennial leaves do fall off after few years.

The longevity of the saxicolous lichens like *Rhizocarpon*, usually growing on rocks in alpine regions have been made use of in assessing the age, or dating the retreating glaciers by employing lichenometric methods. For this, the growth rate of these lichens on the moraines is found out, and the age of the exposed moraines calculated on the basis of the size of the lichen thalli.

#### *Relationship with Bryophytes*

Normally no biological relationship exists between the lichens and the bryophytes except that the latter generally act as substrates. But a small group of lichens, numbering 23 taxa have been found associated with the

bryophytes in a more or less parasitic manner with strong specialisation for the host. In the lichen taxon *Vezdaea aestivalis*, it has been demonstrated that the photobiont *Leptosira obovata* is sustained subcuticularly in living bryophyte leaves for sometime before they die.

### *Pedogenic Activity of Lichens*

Lichens are the only plants which are able to grow on barren rocks and make use of the conditions which are unfavourable to other groups of plants. They have the adaptability to stick, penetrate and continue their slow growth for a long period of time under their unique ability to tide over prolonged desiccation and extremes of climatic conditions. In this process, the weathering of the rocks takes place with the formation of the soil, which is beneficial for the growth of other plants. But the same lichens become a cause of concern and are abhorred when they claim a place on the stone monuments and sculptures, disfiguring them in the process.

### *System of Food Transport in Lichens*

The survival and the wide distribution of the lichens has been possible on account of the almost nonexistent system of food transport. The mycobiont and the photobiont continue to grow almost side by side to the extremities of the thallus, so that the photobiont is always capable of synthesising the food for both the bionts. The vegetative mode of propagation ensures that the photobiont (kitchen) is always associated with mycobiont, be it fragmentation or formation of soredia and isidia. Carrying of the photobiont cells along with the ascospores has also been achieved in some lichen taxa, in which the photobiont cells are present in the epithelial region, e.g. *Endocarpon* sp.

### *Species pairs and selection pressure for evolution*

Poelt (1970) denoted the occurrence of species pairs in several genera of lichens. One of the species of the pair is fertile with ascocarps, and the other is sterile with vegetative propagules. The thalli of the two are morphologically and chemically very similar. Examples of these are: *Parmelia cetrata*/*P. reticulata*; *P. nilgherrensis*/*P. pseudonilgherrensis*; *Physcia ciliata*/*P. orbicularis*, *P. leptalea*/*P. tenella*; *Xanthoria elegans*/*X. sorediata* etc. In taxa which possess ascocarps and also vegetative propagules (soredia or isidia) it has generally been observed that there is often an inverse proportion in the development of the two. When environmental factors favour the development of vegetative propagules,

there is a corresponding decrease in the formation of the ascocarps. It is likely that completely sterile taxa have evolved from similar situations by eliminating the ascocarp formation, or their retention by other factors. The selection pressure has favoured the sterile forms for their survival and wide distribution as they possess more efficient modes of reproduction in comparison to the reproduction by ascospores. This type of progressive sterilization is also seen in higher plants.

### *Fertility of the Photobiont*

In general, the increase of the photobiont cells in the lichen thallus is by simple or mitotic divisions. In very few cases only the formation of the gametangia and asexual spores have been reported. While the sexual reproductive activities of the mycobiont are so prolific, the behaviour of the photobiont is strangely opposite. When the green algal photobionts are cultured in nutrient solutions they reproduce sexually or asexually indicating that they have not lost the power of reproducing that way. No definite reasons have been assigned for a different behaviour of the photobiont in the lichenized state. It is probable that the mycobiont produces some sort of enzyme or hormone which inhibits the sexual reproductive activity of the photobiont. It has been reported that when a sporangiophore-producing *Cephaleuros* colony growing on leaves is lichenized by *Strigula*, the production of sporangiophores is depressed and finally ceases. The photobiont *Phycopeltis* is also reported to behave in the same manner when *Porina* lichenizes it.

### *Lichen Metabolites*

Certain metabolites of excretory nature are formed in the lichen thallus. While some of these can be formed by the mycobiont alone, as evidenced in the pure cultures of the mycobiont, the majority of the metabolites are the products of the lichenization. Since no excretory system is present in the lichens, these metabolites (lichen substances or lichen products as they are also called) get deposited on the surface of the hyphae in different tissues and are therefore also known as extra-cellular constituents of lichens. Certain metabolites occur only in the cortex, others in the region of the medulla or in the tissues of the ascocarp. More than 220 such lichen substances are now reported and their chemical configurations worked out. They are basically grouped into primary products of metabolism, products of acetogenine, phenylalanine derivatives and vitamins. The aromatic compounds of the acetogenine activity predominate and some of them are responsible for imparting the characteristic colours to the lichen thallus or

its parts. For example, the presence of anthraquinone, parietin, in the cortex imparts it an orange-red colour, usnic acid imparts a yellow colour to the cortex. A large number of depsides and depsidones are usually present in the medulla. All the lichen substances are almost uniformly distributed in the region of the thallus in which they occur. A small fragment of the lichen thallus or its particular part is sufficient to demonstrate the presence of the particular lichen substances by employing microcrystallographic and thin layer chromatographic methods. Both the methods are elaborately much advanced as far as the determination of the lichen substances is concerned. Several of them give characteristic colour reactions with potassium hydroxide and calcium hypochlorite solutions in water and paraphenylene-diamine in alcohol. Colour tests help in indicating certain groups of lichen substances and are extensively used for taxonomic determination of lichen taxa.

The lichen substances though referred to as excretory, render an enviable advantage to the lichen thalli. They often act as antibiotics and protect the lichen thallus from microbial attacks. Due to a bitter taste, lichens are not usually eaten directly as fodder. Pigments increase or reduce intensity of light. Air dried lichen specimens in the herbarium usually keep their colour indefinitely and hardly need any preservative.

#### *Indicators of Atmospheric Pollution*

As far as the adaptation and survival under extreme climatic and edaphic conditions are concerned, lichens can be considered quite hardy. But when the percentage of  $\text{SO}_2$  goes up in the atmosphere above a particular minimum lichens are the first plants to be affected and therefore are called indicators of pollution. The  $\text{SO}_2$  causes bleaching of the photobiont cells, resulting in the stoppage of the photosynthesis and subsequent death and disintegration of the lichen. The taxa which are most susceptible are: *Physcia aipolia*, *Anaptychia ciliaris*, species of *Ramalina* and *Usnea*. On the contrary, there are several lichen taxa which are tolerant to the pollution. Detailed work on these aspects has been done in the European countries and lichens have been found useful in monitoring the atmospheric pollution.

#### *Evolutionary success of the lichenization*

The photobionts in general, except probably *Trebouxia*, are fairly well distributed in free state in nature, but the mycobionts are not known to be living in nature in the free state (as parasites and saprophytes). There is

still inadequate knowledge about the exact manner of the development of lichens in nature. It is likely that during the evolutionary progression some fungi could develop a symbiotic relationship with the algae or cyanobacteria in preference to the parasitic-saprophytic mode of living. This ensured the survival of the bionts under extremes of climatic and edaphic conditions. It has been suggested that the ascospores of lichens (mycobiont) on germination may come in contact with any type of algal cells and behave in parasitic manner till the requisite photobiont comes in contact with the hyphae. The remark "Thus lichens clearly indicate how two dissimilar organisms, under unfavourable conditions, can survive by forming a cooperative union, an example which man could well emulate" by Ahmadjian (1988) seems particularly suited to people of present day India.

### ACKNOWLEDGEMENT

I am grateful to the Indian National Science Academy, New Delhi for the award and the honour to deliver this lecture at this ancient city of Allahabad situated at the confluence of the sacred Ganges and the Jamuna rivers.

### References

1. Awasthi D D 1975 Evolutionary trends in lichen symbiosis; in "*Advances in Mycology and Plant Pathology*" pp. 125-130 ed S P Raychaudhuri et al.
2. Ahmadjian V 1982a Definition of the term lichen. *International Lichenological Newsletter* 15 18-19
- 1982b Algal/fungal symbioses. *Progress in Phycological Research* 1 179-233
- 1988 Coevolution in lichens; *Ann. New York Acad Sci* 503 307-315
- and Jacobs J B 1983 Algal-fungal relationship in lichens: recognition, synthesis and development; in "*Algal Symbiosis*" pp. 147-172 ed L J Goff Cambridge University Press
3. Hawksworth D L 1988 The variety of fungal-algal symbioses, their evolutionary significance and the nature of lichen; *Bot. J. Linn Soc.* 96 3-20
4. James P W and Henssen A 1976 The morphological and taxonomic significance of cephalodia; in *Lichenology. Progress and problems* pp. 27-77 ed D H Brown et al (London: Academic Press)
5. Poelt J 1970 Das Konzept der Artenpaare bei der Flechten. *Deutsch. Bot. Ges neue Folge* 4 187-198



**Ramdeo Misra** (b. 26 August 1908) did Ph.D. (1937) from University of Leeds, UK. He is Professor, Head of Botany Department, University of Saugor, Sagar, and Banaras Hindu University (BHU), Dean of Faculty, and Vice-Chairman of Teaching Council, BHU, Varanasi.

Misra has made innovative measurements of soil redox potential and established its value for plant adaptations along soil-moisture gradient. Developed a theory of body- biomass build-up. Demonstrated ecotypic differentiation. Discussed synecological and ecosystem hypotheses in the light of tropical

vegetation of India. Examined ecology, social value system and environmental education.

Misra is Fellow of National Academy of Sciences (India) and Indian Botanical Society (President, 1959); Founder Fellow, National Institute of Ecology, World Academy of Arts and Science and International Society for Tropical Ecology (President, 1971-75); he has been Member of INSA Council (1981-83). He was President of Botany Section, Indian Science Congress (1958); and is recipient of Birbal Sahni Gold Medal (1967); Jawaharlal Nehru Gold Medal (Madhya Pradesh Vigyan Academy) (1974); Bronze Medal International Society for Tropical Ecology (1979); Sanjay Gandhi Award for Environment and Ecology (Government of India) (1984), Professor T.S. Sadasivan Endowment Lecture (INSA) (1984); Swami Pranavanand Saraswati Award (UGC) (1986)

Commemorative medals have been instituted in honour of Ramdeo Misra, e.g. Founder's Medal by International Society for Tropical Ecology (instituted in 1979); Ramdeo Misra Gold Medal by Indian Environmental Society (instituted in 1980).

*Ramdeo Misra was elected to the fellowship of the Academy in 1955*

**Address:** B-7, Indian Agricultural Research Institute, New Delhi 110 012;  
**Tel:** (011) 5745102.

## **UNIVERSITY EDUCATION IN ENVIRONMENTAL SCIENCE**

R MISRA

It is a great pleasure and privilege to deliver Professor T S Sadasivan Endowment Lecture under the auspices of INSA-Varanasi Chapter. Professor Sadasivan is known for his deep insight into the ecology of host-parasite relationships. He was instrumental in initiating the University Grants Commission in 1971 for the institution of the Advanced Centre of Studies in Ecology at Botany Department of Banaras Hindu University where I have taught and learnt the subject for about 40 years. Hence I am indebted to INSA and Professor D P Burma, the convener of the local Chapter respectively for awarding me the Lectureship and arranging the lecture at my 'alma mater'.

Few will disagree with the need of environmental education as aid for reorientation and support of our life style. Nevertheless, much deeper philosophical questions to which complete answers are not available at the moment, are involved in making the decision about the scope and content of such education. Our divergent views on the subject and intra-and inter-departmental rivalries are responsible for stalling action, especially at the University level. What kind of expertise, understanding and behaviour are expected of scholars and tomorrow's leaders in a society beset with consumerism in leading to urbanisation and industrialisation in an era of ecological scarcity? An attempt is being made to meet this challenge through relevant and effective environmental education.

Society and environment are always in transition. Political, administrative and legal institutions are created to make the transition as smooth as possible. Education particularly in the Universities provides leadership for the purpose.

Environment is not only a matter of perception but an objective reality in which we are embedded for functioning. In turn the way we develop values and attitudes, interacts, and impacts, modifies the environment including ourselves to the extent of the experienced stresses. Thus environment is a web of perceived systems and factors which we can indicate, qualify and/or quantify for manipulating it in order to sustain a more spiritually and materially satisfying life. We use ecology as tool for

understanding the environment as a system. Ecology is growing rapidly by internalising the different disciplines of natural and more recently social sciences. These other disciplines have also gained from ecology by applying many of its principles to their own discipline. Since ecology is rooted within the life sciences, interdisciplinary rivalries in the university system have made many scholars blind to its fertility in promoting a healthy growth of their own sciences. Indeed no discipline has ever progressed without seeking information from other parts of knowledge as the latter is indivisible in the ultimate analysis.

When we encounter human environment and develop human ecology on the principles based on life sciences, a reorientation of knowledge from all the sources is demanded. The skill for doing so makes ecology in the broadest sense like the skills we develop through education within each of the disciplines traditional to the university. These traditions have been zealously guarded according to predisposition of the scholars and the limitation of time devoted to training. Nevertheless every now and then universities come up with proposals for instituting hybrid departments such as biochemistry, biophysics, econometrics, business management, etc. etc. thus breaking the barriers among the chosen disciplines.

Realising the importance of ecology and newer avenues of expanding knowledge traditional disciplines have incorporated in their studies environmentally oriented courses such as environmental biology, environmental physics, chemistry, geology, law, politics, engineering, etc. etc. In fact all the 120 teaching departments of the Banaras Hindu University interface with the environment in their respective fields. So a host of environmental courses are being administered. These activities though sectoral are highly desirable and as important as the schools of environmental sciences established at the Jawaharlal Nehru University, Andhra, Kerala (agriculture), Annamalai Universities etc. trying to understand the impact of specialised disciplines on the environment. Suitable degree and diploma courses as well as those related to in service and public administration oriented courses are designed for appreciation and reorientation in matters related to the environment and its enhancement. Nevertheless, they do not impart in-depth training in the management of ecosystems. Managers are required to foresee the impact of human activities on the environment from holistic approach and even advise specialists where they worsen the situation on account of their fragmentary knowledge and zeal for improving the environment. Very



often well intentioned steps to reduce pollution, improve irrigation or drainage, plan human settlements, etc. etc. become counter productive because either the scale or the holistic approach of the ecologist has not been taken into account. Slogan shouting and agitations raised by the public are also focussed on political and so called ecological issues which are not examined in the context of the whole resulting into disproportioned growth spawning more and more difficult problems of the environment, since these are devoid of perspective ecological planning on a suitable scale of time and space. Hence, we need leaders and managers for environmental management. This task needs a thorough grounding in ecology and relating social trends with the way resources are obtained, used and consumed.

In view of the above discussion it is proposed to install a full fledged two year M.Sc. course in environmental science which has to be erected as a single discipline training the mind to simulate environmental systems by transcending related traditional disciplines. Ecology has opened the door for such graduation. The subject oriented diploma and short term fragmentary courses are not relevant to either policy making or environmental designing for improving or maintaining the dynamic human environment as a whole.

Obviously environmental science as a single discipline cannot be the sum of all the environmental sciences including arts, physical and biological sciences, social sciences, etc. It ought to train the mind to establish linkages of interactions among them as observed in man's real world.

The real world can be observed as a set of phenomena in three distinct ways: (i) as discrete events, (ii) as sets of repetitive phenomena, similar individuals or populations or even aggregates of dissimilar populations, and (iii) as a set of interlinked processes or system. It is the systemic and holistic approach of ecology which differentiates it from other disciplines based on individuals and populations. Thus environmental education is training of the mind to grasp meaningful dynamics of the system of human societies and environments in transition as a whole. A person so trained can at once link up human activities with his surroundings extending right up to the biosphere. By quantifying the parameters one can develop a predictive model in which detailed processes can be telescoped or transcended according to the object in view such as impact analysis. A system model of any scale can be qualified and

quantified, described in words or depicted in flow diagrams besides developing mathematical models.

The holistic study demands consideration of the functioning of the whole in relation to the subsystems within it. The meaning of the human environment thus unfolds itself within the context of a hierarchy of systems functioning within the biosphere and the universe. The functioning is achieved through the great biogeochemical cycles from which life has been evolving. The five basics of the cycle remain soil (earth), water, energy, space and air as conceived by the Indians from the Vedic time. The combinants, breakdown and recombinations of elements of these within space exhibit always newer phenomena not possessed by the combinants. The 'Brahmāṇḍa' or the Universe is the largest system containing all the real world phenomena. By subsuming the comprehensible systems as subsystems or ecosystems and with the aid of analytical and synthetic exercises we strike at man and his societies as evolving subsystems of the ecosystem of a given environment.

Man is a comparatively recent phenomenon of the biosphere. Besides his physical existence he is conscious of his perceptions of psychosocial, political and economic organizations, centred around the resources of living. He has also developed awareness of yet higher values of life such as spiritual, aesthetic and moral—very often subsumed within religion. These values interact and he expresses them in guiding his conduct through a hierarchy of value systems. Individuals, families and societies are guided by such orientation within the environment and so environment and society ever remain in transition. Thus environmental education has no meaning without an appreciation of the multidimensional man who has become the agent of change in both.

In the past era of ecological abundance science and technology propelled human institutions to exploit the resources of the biosphere. During the past two centuries industrial growth has exponentially accelerated the process. Now we are facing the stark reality of ecological scarcity. Nation states had differentiated during the period of ecological abundance. Those having precedence in resource exploitation have become disproportionately rich and so we have economic disorder leading to confrontation among the North and the South and the West and the East. Meanwhile ecological scarcity is compounded with two side growths of science and technology. These are nuclear power and communication explosion. So the world has become much smaller shivering with the

prospect of nuclear holocaust and a much worse ecological scarcity with the denial of social justice. This challenge can be met only by a deeper appreciation, understanding and appropriate action through environmental education using all the means of communication.

It is naive not to recognise the pivotal role of environmental education for bringing about transformation in our political, legal and administrative institutions designed to run smoothly societies right from village Panchayat to the United Nations Organisations for the wise use of systems of environmental resources. We must realise that military threats to our survival are becoming more and more absurd. Never before was mankind capable of destroying itself not only as a possible result of the world-wide arms race but also as a result of the uncontrolled exploitation and destruction of the global resource base. The situation can be averted only by tempering our ego and greed through proper environmental education. It cannot be achieved by technological tinkering of the problems of pollution or settlements. We have to redesign a conservation oriented post-industrial society based on innovative and spiritual values.

In view of the scope and content of environmental education and reasons thereof given above the following suggestions are made for a proposed two year M.Sc. course:

#### *Two Year M Sc. Degree Syllabus*

### ENVIRONMENTAL SCIENCE

1. Comprehension of the holistic dimension of the environment and man's place in it. The ecosystem as a web of dynamic relationships.
2. Basic ecological principles. Function, structure and evolution of populations and ecosystems. Driving and state variables of ecosystems. Limits to growth. Growth and Development. Density stresses.
3. Natural, man modified and built environments. Resources, their base and use systems. Wastes and pollution. Industrial growth.
4. Public and private enterprises and ecological scarcity. Governmental and social organisations. Global and international institutions concerned with resource management. Depletion of the environment. War and peace.

5. Competitive uses of land, water and air. Environmental problems. Their origin, intensification and sustainable solutions. Conservation, cooperation and coordination for resources use within ecosystems. Tragedy of the commons and the value of sacrifice and love for all beings.
6. Ethics, aesthetics, culture, spiritual and moral values of life. Place of coercion and education. Value reorientation and environment. Mechanism of social changes.
7. Redesigning society and environment on variable scales. Modelling of ecosystems and simulation exercises.
8. Expectations from the post graduates in the environmental science. Their role in dissemination of knowledge and skill for environmental management. Public awareness, training and education both formal and non-formal about the state of the environment and planning.

Thanking you; I close.



**Brij Mohan Johri** (b. 11 September 1909) did D.Sc. (1936) from Agra University. He was Hony Professor (1974-80), Retired Scientist (UGC) (1975-77), Head of Department of Botany and Director, Centre of Advanced Study in Botany (1966-73), Dean, Faculty of Science and Executive Councillor (1969-70), at University of Delhi, Delhi.

To Johri's credit stands the first and only report of the entry (by suction?) of pollen grains in the style and ovary of *Butomopsis lanceolata*; a truly gymnospermous character in a typical angiosperm Johri reported, for the first time, the largest synergid

haustoria and a 3-nucleate antipodal (cells do not organize) branched haustoria in *Quinchamalium* of Santalaceae. In Loranthaceae the tip of the embryo sac, he discovered, grows up the stigmatic epidermis in *Helianthera* and in *Moquiniella* the tip curves backward and the egg apparatus acquires an inverted polarity. Also found that the outline of the vase-like naked mature endosperm depends on the vascular skeleton of the ovary. Demonstrated, for the first time, the growth responses of the mature endosperm — long considered to be a dead tissue — of *Exocarpus cupresiformis* in aseptic cultures, and that the endosperm and embryo of semi-parasites (Santalaceae and Loranthaceae) could be cultured without any physical contact with the host tissue or the addition of its extract to the nutrient medium.

Johri is Fellow of Indian Botanical Society (President, 1965), Founder Member, International Society of Plant Morphologists (President, 1966-69), International Association of Plant Tissue Culture, Plant Tissue Culture Association (India), (Indian) Society of Plant Taxonomy, Asian Association of Biology Education, Indian Association of Biological Sciences (President, 1966-80) and (Indian) Society of Environmental Scientists. He is the recipient of Birbal Sahni Gold Medal (1970); Professor T.S. Sadasivan Endowment Lecture Award (INSA) (1986).

*Brij Mohan Johri was elected to the fellowship of the Academy in 1956.*

**Address:** Professor (Retired), University of Delhi, Delhi 110 007; Res: Riviera Apartments, Flat 29N, 6th Floor, 45 Mall Road, Delhi 110 054; Tel: (011) 2527695.

## PERSPECTIVES IN ANGIOSPERM EMBRYOLOGY

B.M. JOHRI

*The reproductive biology of angiosperms has posed many questions to which attention should be paid to determine the regulatory and control mechanisms at the macro- and microlevels. Some of the questions discussed in this paper relate to the (a) occurrence of embryo sac-like structures in anthers, (b) number and disposition of nuclei in a female tetrad, dyad, coenomegaspore, and mono-, bi- and tetrasporic embryo sac, (c) formation of Nuclear, Helobial and Cellular type of endosperm, and (d) highly specialized mode of development of embryo sac and endosperm in Loranthaceae, embryo sac in Santalaceae, nutrition of embryo sac in Lentibulariaceae, and development of embryo in Cuscuta and Paeonia*

### INTRODUCTION

The embryology of angiosperms has been studied for almost 150 years, and the Maheshwari School has been the most active and important international centre. This centre attracted a number of foreign botanists to work with us on *in vivo* and *in vitro* systems. For sustained work on seed plants, Professor P. Maheshwari was elected a Fellow of the Royal Society of London — possibly the only embryologist FRS.

Maheshwari's pupils at the Agra College, Agra (1931-1936) and at the University of Delhi (1949-1966), the second and third generation pupils spread all over India, and some abroad, have significantly added to our knowledge of reproductive biology of flowering plants; more than 60 per cent (arbitrary estimate) still remain to be studied. To visualise that no new knowledge will emerge from the study of uninvestigated plants is a fallacy.

### EMBRYO SACS (♀) IN ANTHERS (♂)

Němec (1898) discovered that the pollen grains in petaloid anthers of *Hyacinthus orientalis* are of two types, smaller ones devoid of reserve food, and larger ones packed with food reserve. In the latter the generative cell degenerates. On germination the protoplast enlarges and elongates and the vegetative (tube) nucleus divides. The two- and four-nucleate pollen becomes polarized and simulates embryo sacs. With the next division an eight-nucleate pollen-embryo sac is formed (Němec phenomenon). The

pollen exine remains stuck at the base of this embryo sac. The pollen-embryo sacs are mostly associated with 90% collapsed pollen grains.

Florists keep the bulbs indoors in subdued light at 50°F for two weeks. Then the bulbs are subjected to 70°F. The temperature treatment leads to the formation of pollen-embryo sacs (see Johri & Ambegaokar 1984).

My former student Manasi Ram (1959a) discovered an even more interesting phenomenon in *Leptomeria billardieri*. While normal pollen grains are formed as usual, sometimes one or more sporogenous cells in the anther enlarge, the nucleus undergoes three successive mitotic divisions, from an early stage the nuclei become polarized and two-, four- and eight-nucleate structures are formed (figure 1). These simulate the embryo sac with the usual organization – three-celled egg apparatus, two polar nuclei, and three antipodal cells.

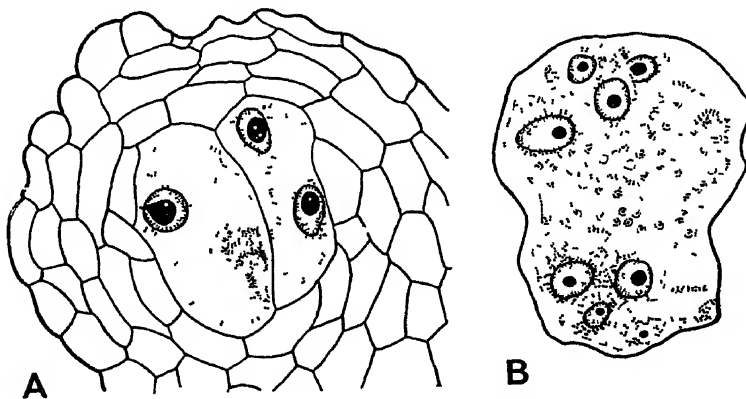


FIG 1 *Leptomeria billardieri*. A Transection anther lobe with two sporogenous cells, one is a binucleate 'embryo sac' B Organized eight-nucleate 'embryo sac' (After Ram 1959a)

Němec (1898) pointed out that the pollen grain acquires 'female potency' and develops into 'embryo sac'. Stow (1934) also presumed that under some abnormal conditions the dominant 'male potency' of pollen grain is weakened or disappears, the 'female potency' expresses itself leading to the formation of pollen-embryo sac. A decisive investigation to demonstrate 'male' and 'female potency' in pollen grains has yet to be programmed. Whether there is any possibility of the control being exercised by male and female sex hormones (= male and female potency mentioned above) should be thoroughly investigated so that this concept can be proved or disproved. In view of the fact that the occurrence of pollen-embryo sacs is so rare in angiosperms, the sex hormone theory appears to be totally unconvincing.

## EMBRYO SACS IN OVULES

The number of meiotic and mitotic nuclear divisions and the disposition of nuclei lead to the formation of a megaspore tetrad, dyad and coenomegaspore which give rise to various types of embryo sacs (see Johri and Ambegaokar 1984). A monosporic bipolar eight-nucleate embryo sac derived from the basal megaspore of a tetrad occurs in *Polygonum*. In *Balanophora* the egg apparatus is always organized at the chalazal pole, instead of the micropylar pole. In *Oenothera* the monopolar four-nucleate embryo sac (antipodals and one polar absent) develops from the micropylar megaspore (figure 2).

A bisporic bipolar eight-nucleate gametophyte develops from the upper dyad in *Endymion*, and from the lower dyad in *Allium* (figure 2).

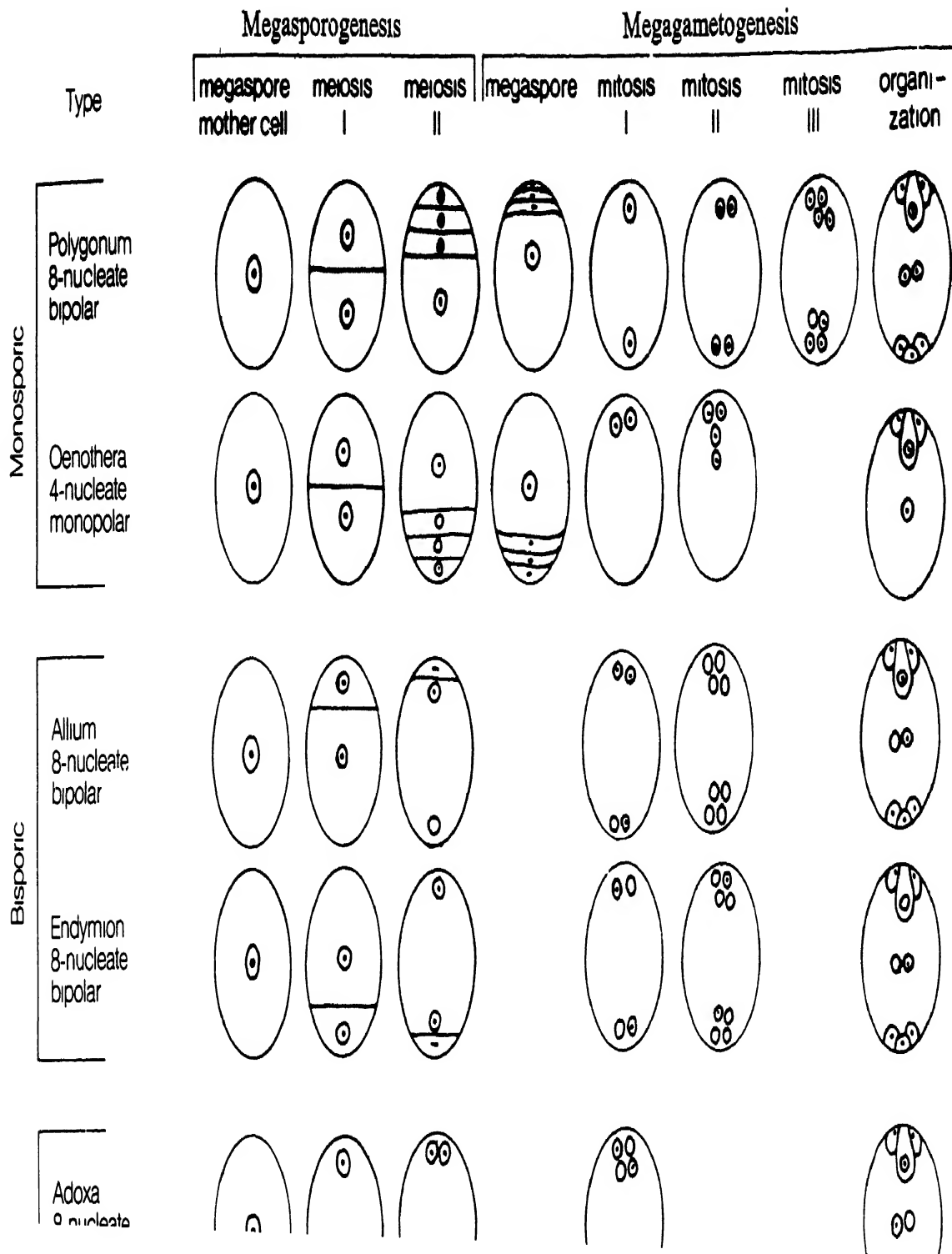
All the four megaspore nuclei participate to give rise to tetrasporic bipolar eight-nucleate embryo sac in *Adoxa*, tetrapolar 16-nucleate in *Penaea*, tetrapolar eight-nucleate in *Plumbago*, polypolar 16-nucleate in *Peperomia*, bipolar 10-(3 polars, 4 antipodals) and 12-nucleate (7 antipodals) in *Chrysanthemum*, bipolar 16-nucleate (11 antipodals) in *Drusa*, bipolar eight-nucleate (1 polar and 3 antipodals triploid) in *Fritillaria*, and bipolar four-nucleate (1 polar and the sole antipodal triploid) in *Plumbagella* (figure 2).

Under each type many variations occur in the development and organization of the embryo sac; in the same taxon more than one type of embryo sac also occurs. For example, in different species of *Tamarix* (see Johri and Kak 1954) four types of tetrasporic embryo sacs occur: *Fritillaria* type 38-90%, *Drusa* 3-10%, *Adoxa* 2-48% and *Chrysanthemum* 6-47%.

What we ought to investigate is what mechanisms are involved (a) which prevent wall formation after meiosis I (bisporic) and after meiosis II (tetrasporic gametophytes), (b) which control the number of mitotic divisions and polarity of daughter nuclei resulting in 13 types of embryo sacs, (c) which bring about occurrence of more than one type of embryo sac in the same genus—mono-, bi- and tetrasporic as in *Erigeron* (figure 3), and one or more type of tetrasporic embryo sacs as in *Tamarix*, and (d) whether temperature, humidity and day-length affect the developmental type and organization of the embryo sac? Some reports do indicate that the environmental factors do affect the development. However, the data is insufficient to reach any conclusion.

FIG 2 Schematic representation to show the origin and development of various types of embryo sacs. (After Willemse and van Went 1984) →





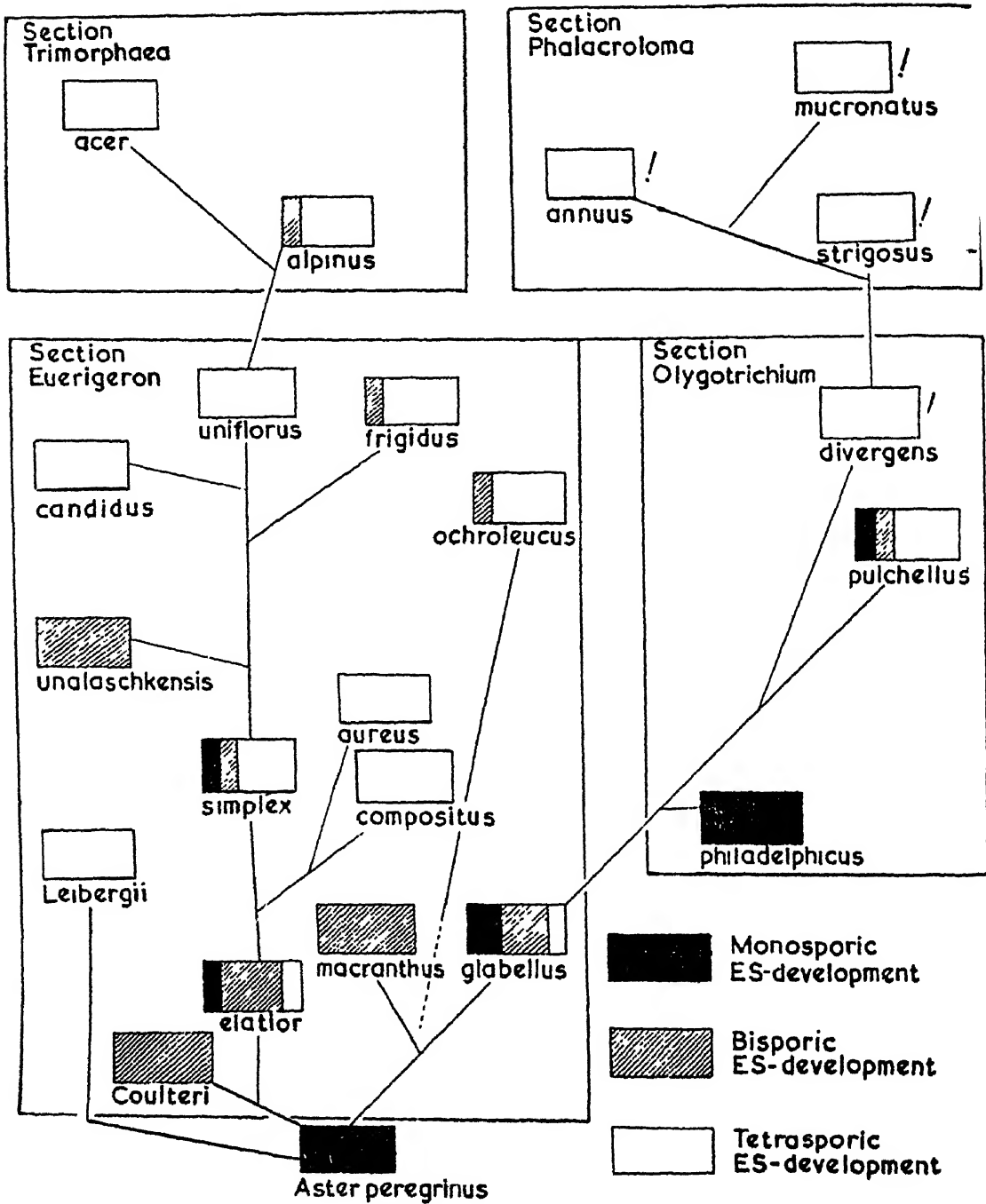


FIG 3 *Erigeron*, occurrence of mono-, bi- and tetrasporic embryo sacs. The exclamation mark (!) indicates agamosperous species. (After Harling 1951)

## ENDOSPERM

The primary endosperm nucleus divides, without any wall formation, followed by free-nuclear divisions (Nuclear type), or a wall may separate a large micropylar and a small chalazal chamber (Helobial type), or wall formation occurs after each division (Cellular type) (figure 4). A wide range of endosperm haustoria is reported in different plants (Vijayaraghavan & Prabhakar 1984). The endosperm is a nutritive tissue.

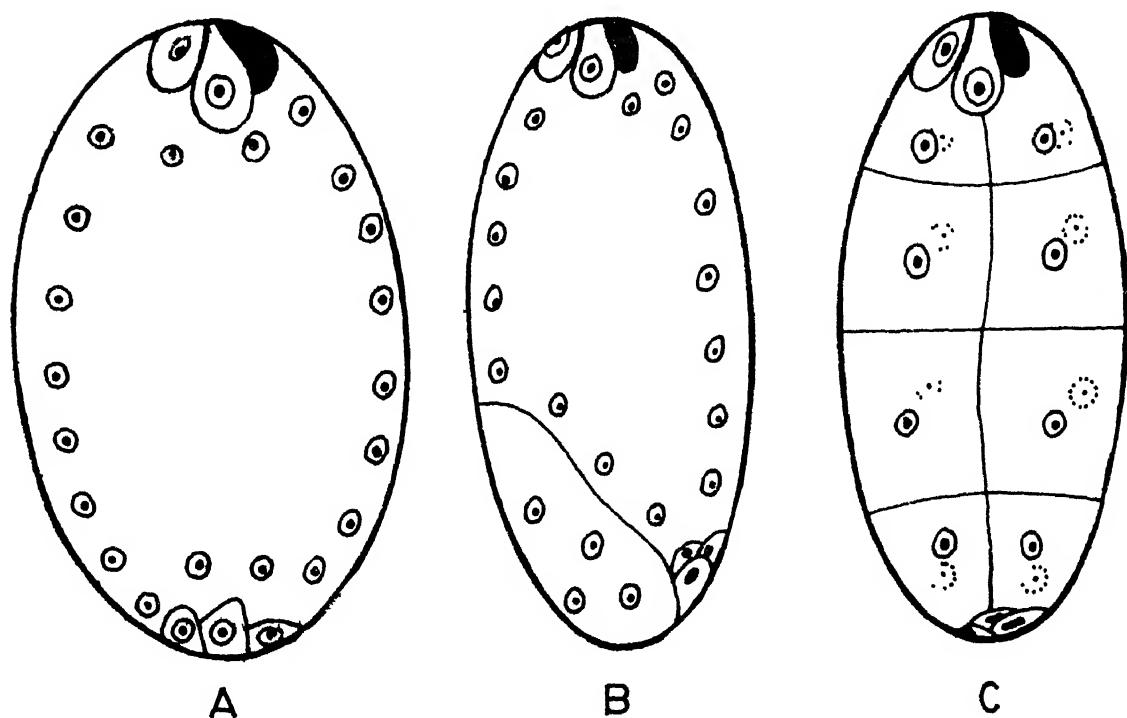


FIG 4 Endosperm types A Nuclear, B Helobial-micropylar chamber has many free nuclei, chalazal chamber has four nuclei, C Cellular type. Note one synergid, zygote, remnants of pollen tube (black) and three antipodals in A, B and C (After Erdelská 1981)

In Orchidaceae, the endosperm, when present, is of the Nuclear type and only 12 nuclei are reported in *Vanilla* and 16 in *Galeola* (Chugai 1971). In Trapaceae (Ram 1956) the primary endosperm nucleus persists up to the formation of a several-celled pro-embryo. In *Trapa natans* (see Ishikawa 1918) it divides once or twice but the two or four free nuclei degenerate. The members of Podostemaceae lack triple fusion and the primary endosperm nucleus is not formed (Mukadda 1962). In this family, in the absence of endosperm, the developing embryo derives nutrition from the pseudo-embryo sac and the haustorial basal cell of the pro-embryo. In the Trapaceae and Orchidaceae the haustorial embryonal suspensor provides nutrition.

In *Butomopsis* (Johri 1935, 1936) the lower polar nucleus is not formed. One of the sperm cells fertilizes the egg, the other male gamete fertilizes the upper polar nucleus. Thus, the embryo and Helobial endosperm are both diploid and genetically identical (both the sperms are sister cells, as also the egg and upper polar nucleus).

Krishnamurthy (1988) emphasises that the endosperm regulates the pattern of development (and the differentiation) of the embryo. The Orobanchaceae (Tiagi 1951) have a well-developed endosperm but the embryo remains globose and undifferentiated (Kuijt 1969). Even in the absence of the endosperm, in Podostemaceae the dicotyledonous embryo is fully differentiated and organized. Therefore, Krishnamurthy's hypothesis should be further examined.

Concerning *Butomopsis*, we must find out the mechanism which determines the development of embryo and endosperm (also diploid) from genetically identical zygote and fertilized upper polar nucleus.

It has been suggested that when the embryo sac is broad the endosperm is free-nuclear but when the embryo sac is narrow the endosperm is cellular. To some extent this is true but we must look for other demonstratable regulatory mechanisms to explain the development of various types of endosperm.

## EMBRYO

The pattern of wall formation in a quadrant and an octant pro-embryo and subsequent cellularization and organization leads to the development of six distinct types of embryogeny (figure 5). The suspensor also generates various types of haustoria (Natesh & Rau 1984).

When the zygote divides by a transverse wall followed by a vertical wall, the embryogeny conforms to the Onagrad and Asterad types. If the second division is also transverse, the embryogeny is of the Solanad, Chenopodiad and Caryophyllad types. The third and subsequent divisions and organization of the pro-embryo varies from type to type. The longitudinal division of the zygote is not very common and the embryogeny is of the Piperad type (figure 5).

Several pertinent questions concerning the embryos in angiosperms have been raised from time to time. Besides the normal embryos with one cotyledon in monocots, and two in dicots, there are pseudomonocotyledonous embryos with one cotyledon arrested as in *Trapa*

| Type         | Division I | Division II | Division III | Division IV |
|--------------|------------|-------------|--------------|-------------|
| Onagrad      |            |             |              |             |
| Asterad      |            |             |              |             |
| Solanad      |            |             |              |             |
| Chenopodiad  |            |             |              |             |
| Caryophyllad |            |             |              |             |
| Piperad      |            |             |              |             |

FIG 5 Schematic representation of main types of embryogeny, *ca* apical cell, *cb* basal cell; *cc*, *cd* daughter cells of *ca*; *m*, *ci* daughter cells of *cb*; *d*, *f* daughter cells of *m*; *n*, *n'* daughter cells of *ci*; *o*, *p* daughter cells of *n'*; *q* quadrant, *l*, *l'* octant — derivatives of *ca*. (After Natesh and Rau 1984)

(Ram 1956), both cotyledons apparently 'fused' (closely appressed along the inner margin) as in some Loranthaceae (Dixit 1961), completely undifferentiated and organless globular embryo in *Orobanche* (Tiagi 1951), other members of Orobanchaceae, Orchidaceae, and a number of other taxa (see Natesh and Rau 1984).

The causes leading to such significant variations are not understood, and techniques and procedures should be developed to unravel the control mechanisms.

There is also a long standing controversy regarding the terminal placement of shoot apex in the embryo of dicots and lateral in monocots. Swamy and Krishnamurthy (1977) have put forward the view that in monocots too the shoot apex is terminal as in dicots. The shoot apex in monocots need not always be terminal. This aspect requires further comparative developmental study and quantitative data to decide as to what percentage of monocots have terminal shoot apex, and what percentage have lateral shoot apex.

### SPECIALIZED EMBRYO SACS, ENDOSPERM AND EMBRYO

A number of taxa have developed specialized mode of nutrition. Some examples are cited from the semiparasitic families Loranthaceae and Santalaceae, total parasitic family Cuscutaceae, and Lentibulariaceae which includes the insectivorous taxon *Utricularia*.

**LORANTHACEAE:** Among the embryological contributions made by Johri and his associates, the account dealing with the Loranthaceae may be considered as very significant. The development of embryo sac in the Loranthaceae is unique. The members of this family lack normal ovules; the embryo sacs differentiate in the ovary, at the base of the stylar canal, or in the mamelon which is a projection from the base of the stylar canal. Several embryo sacs develop concurrently in an ovary. In a number of taxa, at the 4-nucleate stage, the tips of the embryo sacs elongate into the stylar canal and reach up to various heights. In *Helixanthera ligustrina*

FIG 6 A *Atkinsonia ligustrina*, longisection of ovary with three mature embryo sacs B *Amyema muqueli*, embryo sac up with egg apparatus and upper polar nucleus, at two-thirds the height of style C *Helixanthera ligustrina*, longisection of stigma with upper part of five embryo sacs, the tip of one embryo sac touches the stigmatic epidermis. One embryo sac (left) has a bicelled pro-embryo, others with zygote, two embryo sacs with primary endosperm nucleus D *Monquiniella rubra*, longisection of stigma with four embryo sacs (es), three reach up to the tip of stigma and curve backwards so that the egg apparatus/pro-embryo acquires reversed polarity. E *Nuytsia floribunda*, tip of embryo sac (in stylar region) with lateral caecum. F. Longisection of carpel The names of various taxa indicate the height to which the tip of embryo sac reaches in the ovary, style and stigma. (See John & Ambegaokar 1984)

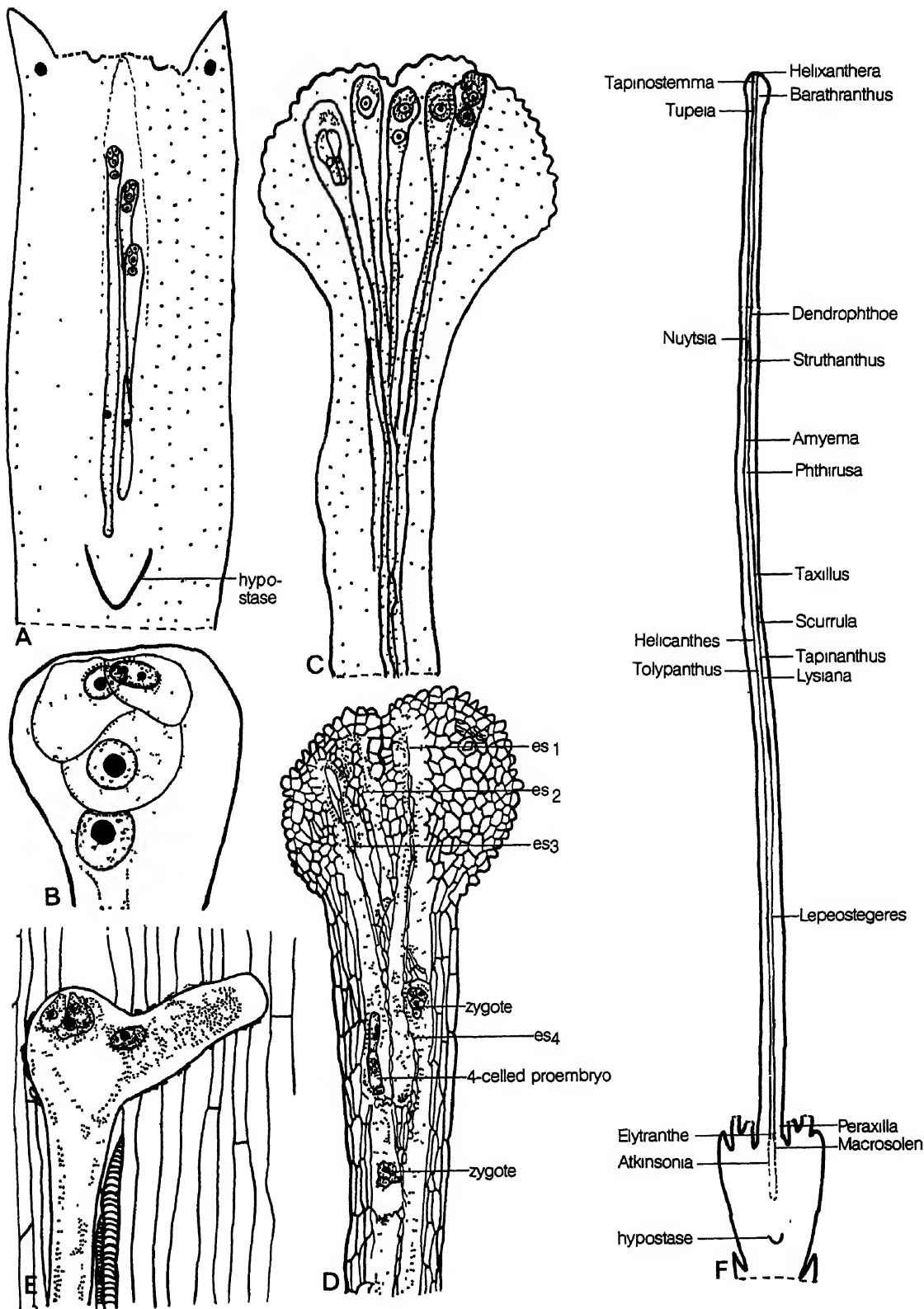


FIG 6

(Maheshwari & Johri 1950) the tip of the embryo sac sometimes grows as far as the stigmatic papillae. In *Moquiniella rubra* 4-9 embryo sacs develop through the 42-48 mm long style and reach up to the base of the stigma (Johri & Raj 1969). Subsequently, the tips of some of the embryo sacs curve backward for 2-4 mm so that the gametophyte becomes — } — or { — shaped. Such embryo sacs have not been reported in any other angiosperm (figure 6).

The endosperm development in the Loranthaceae is also exceptional (Dixit 1961). The primary endosperm nucleus of each embryo sac descends from the style to the lower part of the ovary and divides to form a cellular mass. The individual endosperm proliferates causing the intervening ovarian tissues to become obliterated, and all the endosperms (in an ovary) fuse to form a 'composite endosperm' (figure 7).

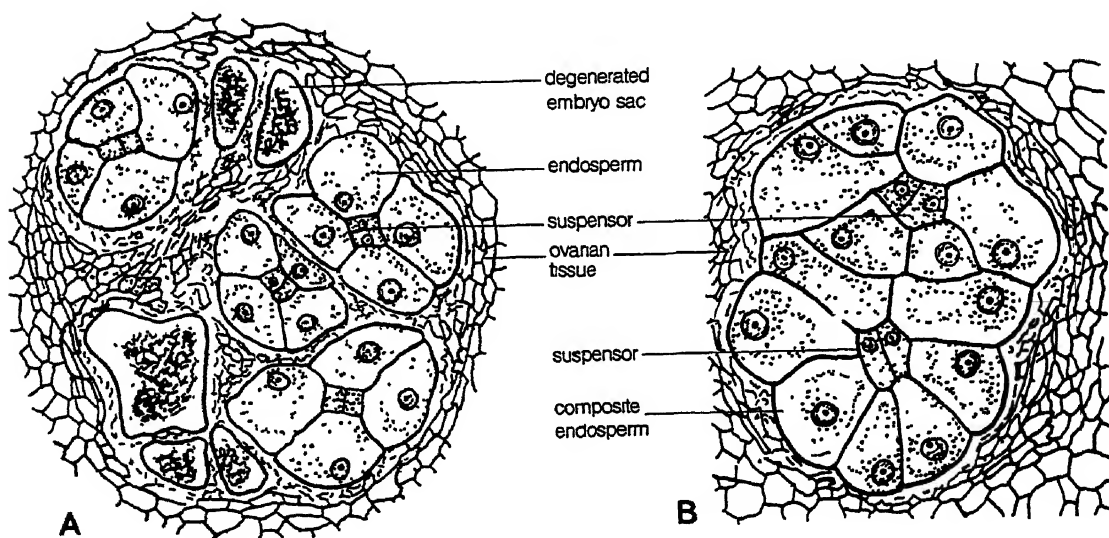


FIG 7 *Tolypanthus involucreatus* A. Transverse section of ovary (central part) shows four embryo sacs, each with four-senate endosperm enclosing bisenate suspensor. Also note five degenerated embryo sacs. B. Composite endosperm formed by the fusion of all the endosperms in an ovary. Note the bisenate suspensor of two proembryos. (After Dixit 1961)

In *Tapinostemma acaciae* Johri & Prakash (1965) recorded the formation of finger-like processes from the basal part of the endosperm. These processes invade the adjoining tissue, sometimes making their way deep into the thick-walled hypostase. This feature of endosperm is not met with in any other member of the Loranthaceae.

**SANTALACEAE:** Johri and his students made a detailed study of the embryology and taxonomy of the Santalaceae. In *Commandra umbellata*



a caecum arises laterally on the funicular side of the gametophyte at the level of the egg apparatus, and grows beyond the ovule into the placenta (Ram 1957). Similarly, in *Leptomeria cunninghamii* the tip of the mature embryo sac grows beyond the ovule, bends upward and extends up to the base of the style. In the meantime, the lateral caecum develops and, after traversing through the ovule, it bends and enters the placenta (Ram 1959b).

In *Quinchamalium chilense* the embryo sac extends beyond the ovule and comes to lie in the ovarian cavity. The synergid haustoria are exceptionally long and reach up to one-third the height of style, attaining a length of up to 1,200  $\mu\text{m}$  before fertilization. The remnants of the haustoria persist up to the globular stage of pro-embryo (Johri & Agarwal 1965). Synergid haustoria of such dimensions have not been reported in any other angiosperm.

CUSCUTACEAE: Most of the species of *Cuscuta* examined embryologically show three-celled pollen at shedding, and Polygonum type of embryo sac. Contrastingly, *C. reflexa*, studied by Johri & Tiagi (1952), showed two-celled pollen as well as a small proportion of three-celled grains at shedding time. The embryo sac conforms to the Allium type. The most interesting feature in the embryology of *Cuscuta reflexa* (Johri & Tiagi 1952) is the occurrence of two types of suspensor, absence of histogenic differentiation in pro-embryo, presence of a filiform and spirally-coiled embryo without cotyledons, and absence of a root cap (Johri & Tiagi 1952). The suspensor may comprise several vesicular coenocytic cells and different embryos show transitional stages so that in some embryos the suspensor is highly reduced and of uninucleate cells. This feature is not known in any other angiosperm.

LENTIBULARIACEAE: In *Utricularia* (Khan 1954) the ovule is anatropous, unitegmic, tenuinucellar and lacks funicular vascular supply. In *U. flexuosa* the micropyle is not organized and the tip of the embryo sac remains exposed. The pollen tube meets the embryo sac in the ovarian cavity (exogamous mode as compared to porogamous, mesogamous and chalazogamous).

The presence of nutritive tissue in the chalazal region of the ovule and adjacent to the placenta is noteworthy. The embryo sac derives nutrition from these nutritive tissues, and through the endothelium.

In this aquatic floating plant the leaf segments are modified into bladders which have a specialized mechanism to trap insects. The insects provide the proteins.

*PAEONIA* (Paeoniaceae): Yokovlev and Yoffe (1957; see also Yakovlev 1969) discovered that in several species of *Paeonia* the zygote undergoes free-nuclear divisions, the nuclei get arranged peripherally, wall formation and cell divisions occur, and at one or more site/s embryonal buds develop. One of these gives rise to a normal dicotyledonous embryo (figure 8). This is a typical gymnospermous character and has not been observed in any other angiosperm (see Johri & Ambegaokar 1984).

Have the features highlighted above originated due to the heterotropic mode of nutrition, as an adaptive function, or are a result of evolutionary tendencies, or due to all the three?

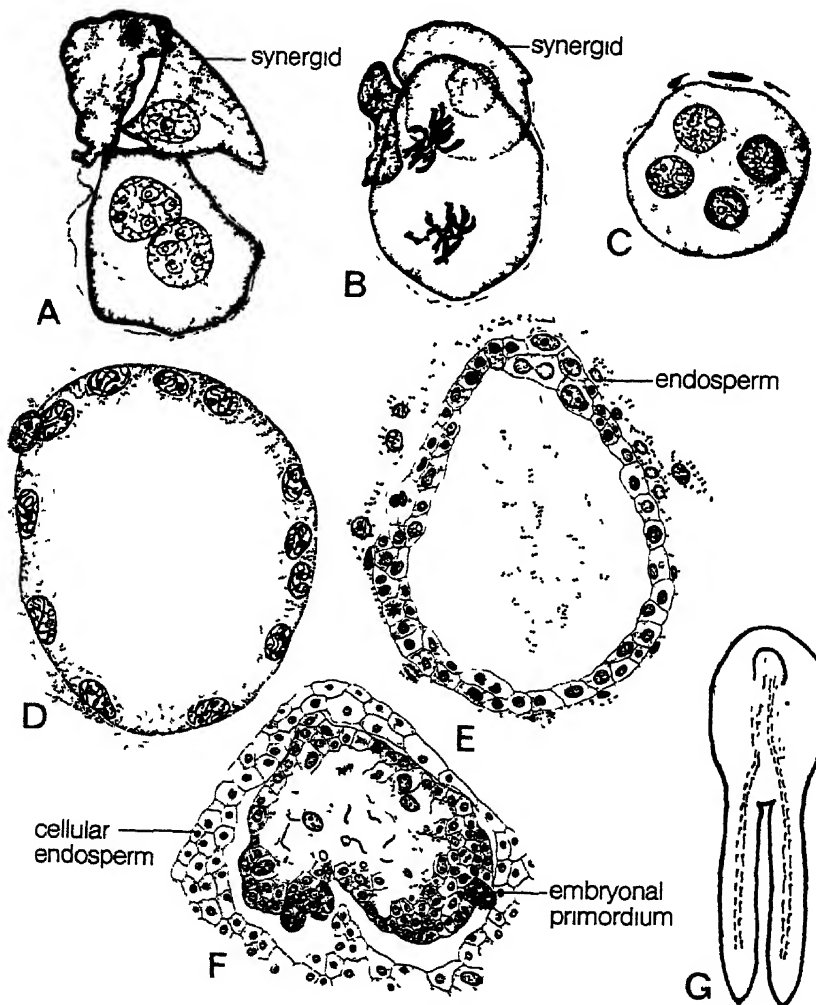


FIG 8 *Paeonia lactiflora* A Two-nucleate pro-embryo. B Both nuclei in division C Four-nucleate pro-embryo D Multinucleate pro-embryo with peripheral arrangement of nuclei E Wall formation in pro-embryo; note some free endosperm nuclei F Differentiation of embryonal buds (primordia) G. Dicotyledonous embryo (After Yakovlev 1969)

## CONCLUDING REMARKS

The questions which have been repeatedly raised are what are the controlling and regulating mechanism/s which bring about the variations in the development of the embryo sacs, endosperm, and embryo? Presumably, the physiological, biochemical and molecular interactions—at the micro-level in the immediate vicinity of the nuclei concerned, at different sites, and at the macro-level in the embryo sac, endosperm and embryo at various stages of development — could affect and control the changes. While we have some information on the ultrastructural and histochemical changes, these are mostly based on fixed material. The information which would be helpful would have to be studied from the live material. However, the probes to undertake investigations on living material have not yet been developed.

A suitable approach could be to add labelled chemicals to the nutrition supplied to the plant (the choice of the plant is very vital), and fix all the developmental stages from flower to fruit (containing mature seeds). Besides cryo-preservation (none of the metabolites would be affected), material should also be fixed in alcoholic (alcohol-soluble metabolites will leach out) and non-alcoholic (only water-soluble metabolites will leach out) fixatives. The sections of all the above materials can be analyzed for the accumulation of labelled compounds in different tissues, the quantitative estimation cannot be made. Coe (1954) fed  $^{14}\text{CO}_2$  to flowering plants of *Zephyranthes drummondii* and made a comparative analysis of accumulation of metabolites (in cryo-preserved and alcohol-acetic acid fixed material) in ovary wall (vascular region and circumjacent parenchyma), ovule (nucellus — micropylar and chalazal region, and integument), and embryo sac (synergids, antipodals, cytoplasm and vacuoles). Coe or any one else has not continued such an interesting investigation.

It is no more possible for a single individual to undertake the investigations to find out the answers to the various questions discussed in this paper. What we need now is a group of people — with competence in different discipline — who should study the various changes and unravel the controlling and regulatory mechanisms which govern the developmental sequence. Probes must be developed to handle live materials so that the approach is to study the dynamic (rather than static) aspects. These problems should deserve highest priority in the twenty-first century.

## ACKNOWLEDGEMENTS

I am very grateful to the Indian National Science Academy for the award of the Professor T.S. Sadasivan Endowment Lecture, 1986, and to Professor T.V. Desikachary for making the arrangements at the University of Madras. Thanks to Ms Binny Mathur for technical assistance.

## References

- 1 Chugai K 1971 Embryological studies of *Galeola septantrionalis* Reichb fil. (Orchidaceae) *Sci Rept Tôhoku Univ Ser. Biol.* **35**: 253-258
- 2 Coe G E 1954 Distribution of Carbon<sup>14</sup> in ovules of *Zephyranthes drummondii* *Bot Gaz*, **115** 342-346
- 3 Dixit S N 1961 Morphological and embryological studies in the family Loranthaceae 8 *Tolypanthus* B1 *Phytomorphology* **11**: 335-345
- 4 Erdelská Olga 1981 Embryologia Krytosemenných Rostlín. Veda, Vydavateľstvo Slovenskej Akadémie Vied, Brauslava (Czechoslovakia)
- 5 Harling G 1951 Embryological studies in the Compositae 2. Anthemidae-Chrysantheminae *Acta Horti Bergiani* **16**: 1-56
- 6 Ishikawa M 1918 Studies in the embryo sac and fertilization in *Oenothera* *Ann Bot (London)* **32** . 219-317
- 7 John B M 1935 Life history of *Butomopsis lanceolata* Kunth. *Nature (London)* **136** . 338
- 8 John B M 1936 The life history of *Butomopsis lanceolata* Kunth *Proc Indian Acad Sci B* **4** 139-162
- 9 John B M and Agarwal Saroj 1965 Morphological and embryological studies in the family Santalaceae 8. *Quinchamalium chilense* Lam *Phytomorphology* **15** 360-372
- 10 John B M and Ambegaokar K B 1984 Embryology : Then and now, pp 1-52. In : John B M (ed) *Embryology of Angiosperms*, pp 1-830. Springer, Berlin Heidelberg New York
- 11 John B M and Kak Dulari 1954 The embryology of *Tamarix* Linn. *Phytomorphology* **4**: 230-247
- 12 John B M and Prakash Sudha 1965 Morphological and embryological studies in the family Loranthaceae 11 *Tapinostemma acaciae* (Zucc.) van Tiegh *Phytomorphology* **15**: 150-158
- 13 John B M and Raj B 1969 Morphological and embryological studies in the family Loranthaceae 12. *Moquiniella rubra* (Spreng f) Balle *Öst Bot Z* **116**: 475-485
- 14 John B M and Tiagi B 1952 Floral morphology and seed formation in *Cuscuta reflexa* Roxb *Phytomorphology* **2**: 162-180

- 15 Khan R 1954 A contribution to the embryology of *Utricularia flexuosa* Vahl *Phytomorphology* 4: 80-117
- 16 Krishnamurthy K V 1988 Endosperm controls symmetry changes in the developing embryos of angiosperms *Proc Indian Acad Sci Plant Sci* 98: 257-259
- 17 Kuijt J 1969 The Biology of Parasitic Flowering Plants Univ Calif Press, Berkeley (U S A)
- 18 Maheshwari P and Johri B M 1950 Development of embryo sac, embryo and endosperm in *Helixanthera ligustrina* (Wall) Dans *Nature* (London) 165: 978-979
- 19 Mukadda A J 1962 Some observations on the embryology of *Dicraea stylosa* Wight, pp 139-145. In: Symp Plant Embryology, pp 1-273 CSIR, New Delhi
- 20 Natesh S and Rau M A 1984 The embryo, pp 377-443 In: Johri B M (ed) Embryology of Angiosperms, pp 1-830 Springer, Berlin Heidelberg New York
- 21 Nemec B 1898 Uber den Pollen der petaloiden Antheren von *Hyacinthus orientalis* L. *Bull Inst Acad Sci Boh* 5: 17-23
- 22 Ram Manasi 1956 Floral morphology and embryology of *Trapa bispinosa* Roxb with a discussion on the systematic position of the genus *Phytomorphology* 6: 312-323
- 23 Ram Manasi 1957 Morphological and embryological studies in the family Santalaceae 1 *Commandra umbellata* (L) Nutt *Phytomorphology* 7: 24-35
- 24 Ram Manasi 1959a Occurrence of embryo sac-like structures in the microsporangia of *Leptomeria billardieri* R Br *Nature* (London) 184: 194-195
- 25 Ram Manasi 1959b Morphological and embryological studies in the family Santalaceae 1 *Leptomeria* R Br *Phytomorphology* 9: 20-33
- 26 Stow I 1934 On the female tendencies of the embryo sac-like giant pollen grains of *Hyacinthus orientalis* *Cytologia* 5: 88-108
- 27 Swamy B G L and Krishnamurthy K V 1977 Certain conceptual aspects of meristems 2 Epiphysis and shoot apex *Phytomorphology* 27: 1-8
- 28 Tiagi B 1951 Studies in the family Orobanchaceae 3. A contribution to the embryology of *Orobanche cernua* Loeffl and *O aegyptiaca* Pers *Phytomorphology* 1: 158-169
- 29 Vijayaraghavan M R and Prabhakar K 1984 The endosperm, pp 319-376 In: Johri B M (ed) Embryology of Angiosperms, pp 1-830 Springer, Berlin Heidelberg New York
- 30 Willemse M T M and Van Went J L 1984 The female gametophyte, pp 159-196. In: Johri B M (ed) Embryology of Angiosperms, pp 1-830 Springer, Berlin Heidelberg New York
- 31 Yakovlev M S 1969 Embryogenesis and some problems of phylogenesis *Rev Cytol Biol Vég* 32: 325-330
- 32 Yakovlev M S and Yoffe M D 1957 On some peculiar features in the embryogeny of *Paeonia* L *Phytomorphology* 7: 74-82



**Chirayathumadom Venkatachaliar Subramanian** (b. 11 August 1924) did Ph.D. (1948) and D.Sc. (1957) from University of Madras. He is Emeritus Scientist (1989- ) and was INSA Senior Scientist (1989-92), Central Institute of Medicinal and Aromatic Plants, Lucknow (1989- )

Subramanian's most significant contributions are on fungal taxonomy, particularly of hypomycetes and their teleomorphs. He proposed a classification of hypomycetes, built on the developmental morphology of anamorph and teleomorph supported by a new terminology. He has done extensive work on soil mycology and soilborne diseases and on soil, coprophilous, litter and other ecological groups of fungi. He has published three books including 'Hyphomycetes: Taxonomy and Biology' (Academic Press, 1983); edited/co-edited two books: 'Taxonomy of Fungi' (University of Madras), 'Mangrove Microbiology' (Unesco, 1988).

Subramanian is Fellow of Indian Academy of Sciences (Vice-President, 1974-85); President, Indian Botanical Society (1972), Mycological Society of India (1985), and International Mycological Association (1977-83); has been Member of INSA Council (1970-71). He is the recipient of Shanti Swarup Bhatnagar Prize (1965); Rafi Ahmad Kidwai Memorial Prize (ICAR) (1972-73); Birbal Sahni Medal (Indian Botanical Society) (1972); President, Botany Section, Indian Science Congress (1975); T.S. Sadasivan Endowment Lecture Award (INSA) (1988); UGC National Lecturer (1972).

*Chirayathumadom Venkatachaliar Subramanian was elected to the fellowship of the Academy in 1960.*

**Address:** Central Institute of Medicinal and Aromatic Plants, Post Bag No. 1, RSM Nagar, P.O., Lucknow 226016; Tel: (0522) 71170, 71568; Res: Anjaneya, Plot 885, 62 Ramaswami Salai, K. K. Nagar, Madras 600078.

## **TRUTH, BEAUTY AND GOODNESS—A PERENNIAL MYCOLOGICAL THEME**

C V SUBRAMANIAN FNA

For the past thirty years and more my research interests have been in mycology and plant pathology, but specially taxonomy and biology of fungi. Any recognition that comes my way is doubtless recognition of the importance the Academy attaches to work in these areas. I am happy about this and I thank the Academy for asking me to deliver this Award lecture.

### **THE BACKGROUND**

I was born in a family which, in a general sense, embraced the legal profession, from my grandfather downwards. I was the first one to go on to a career in science. And this was by accident rather than by design. The family in which I was born had a profound sense of values, ethical and moral, and a profound aesthetic sense which found expression in music. These were coupled with a deep sense of enjoyment of literature and an intense interest in philosophy and historicism, elements that probably retrieved me from the dry and dreary outlook of a scientist bereft of these qualities. That is why I have chosen to speak to you about Truth, Beauty and Goodness, rather than on Science as just Science. After all, the roots of Science are in Natural Philosophy.

If I have been fortunate in my genealogy, I have also been fortunate otherwise. Though the efforts of my English professor in persuading me to opt for an Honours course in English were complimentary, I felt it far better to be able to enjoy literature than to make a formal study of it (how true this is of formal courses in other subjects!). This way I continued to enjoy reading some of the greatest and sublimest in literature available to me. Though I had a preference for chemistry, the College would admit me to an Honours course in Botany and not Chemistry: and, when there is no other choice, rather than being fastidious, we must cultivate what is offered. As a student in the Honours course, I was disarmingly attracted to studying the algae and wanted to continue my study of them, but that also was not to be. Instead, I found myself working on a problem in plant pathology, more precisely in soil mycology. I had a tremendous advantage here. I knew nothing of the subject on which I was to work for the simple

reason that this was not covered in the formal lectures or courses that I had attended. I say 'advantage' because I was entirely left to myself and there was no question of unlearning anything and one of the biggest handicaps one may have to cope with is the effort to unlearn what one may have learnt. It was the same experience when, later in 1950, I went to Kew in England where Mr E W. Mason who was my mentor at the Commonwealth Mycological Institute talked very little, except to say that the fungal specimens and cultures were all open to me and would 'speak' to me, should I want to look at them and benefit thereby! Many specimens carried the incisive and irresistible annotations of Mason in his own handwriting, annotations more eloquent than the man himself. Every fungus that I saw was new to me, a revelation, every fungus a beauty and a marvel, and a treasure, as I learnt as the years went by and I got to know the moulds as one should by their morphology, their behaviour, and their inner power.

I returned to Madras from Kew in 1951 with great enthusiasm and began collecting fungi, chiefly hyphomycetes; their variety, diversity and beauty kept me busy for many years. I did describe a number of them, including some new genera and species, but many more remain undescribed, for identification of fungi is not easy. All this is, no doubt, often considered 'useless' especially, and regrettably, by biologist comrades who bask in the seemingly self-secure and superior feeling of being in line with current fashions and fads in the biological sciences. Not that all descriptive taxonomic work of the kind we are concerned here is of the highest quality, not necessarily, but this is true of all science: there is excellence and there is mediocrity. Mediocrity brings little credit, makes no contribution. All research, all good science, needs to be encouraged and vigorously supported, and it is often more important to support unfashionable areas of research as these usually get crushed for want of support, something that is true of Art and Artists also. Often, this is a case of ideological intolerance. The greatest among our scientists shun fashions and fads and often carry on work in what must be considered unfashionable areas, purely to satisfy the inner urge. Support may not come for such research, but their fascination and passion for research will remain and cannot be suppressed.



## TRUTH, BEAUTY, GOODNESS—THE PERENNIAL MYCOLOGICAL THEME

There is nothing more beautiful than that which is itself unconscious of its beauty. The more the fungi I saw, the more my wonder grew at their extraordinary beauty, and the beautiful things they do. What they are and what they do must necessarily evoke our curiosity. From those who have neither a sense of wonder, nor a sense of curiosity, fungi would still elicit notice from the damage they may do or the benefits they may bestow. Indeed, the present status of the science of mycology is a summation of our appreciation of their beauty, variety and diversity, a reflection of our curiosity and efforts to probe into them to get at the truth about them, our predicament in trying to combat the damage they do, and our hopes of harnessing their power to advantage. Beauty, Goodness and Truth thus form the basic and perennial theme of mycology and the philosophy of the mycologist.

First, about Beauty, Beauty is an experience, our response to Nature, one that fits wonderfully our aesthetic faculties. The perception of Beauty is a prelude to understanding, to knowledge, to Truth. Beauty derives from Nature; it is symbolic of the harmony of form and function that characterises all that constitutes Nature, the harmony of co-existence. As Rabindranath Tagore says, "Man has a fund of emotional energy which is not all occupied with his self-preservation. This surplus seeks its outlet in the creation of Art, for man's civilization is built upon his surplus." The perception of Beauty, and the search for Truth, then, are an expression of the sublimest in Man, of his emotional energy.

About Goodness: Goodness goes with Beauty. The harm that moulds are supposed to do, and the good they are supposed to bring, are both purely our assessment of their 'vices' and 'virtues', but they are both part of their functional beauty. It is good that this is so. We should not subordinate them to ourselves. What with the tremendous advances in technology and the sophisticated and fast life styles that are their aftermath, should not Man at least wonder at the functional beauty of these organisms? Now it is not that he may care to wonder at them, but must, for his own welfare and survival. The moulds have inherited the earth: ours is the mouldy earth.

Now, about Truth; what is Truth? What is Truth in the context of science and of discovery? Truth eludes us and yet we are after it: the progress we make in this our endeavour is in proportion to the intensity of

our curiosity, our passion, to know the Truth, but also in proportion to the quality of our urge especially in framing questions and to the strength of our innovative spirit in using techniques and tools which are themselves a product of progress in science and technology. The result? We go step by step by approximations to Truth, in which elimination of error leading to refinement in ideas and hypotheses or theories plays a significant role.

To illustrate, I should like to share with you the progress we have made in classifying the moulds, in the taxonomy of Hyphomycetes, the fungal group to which most common moulds belong. This can form the subject of a series of lectures, but I must now merely summarise briefly the essence of the situation.

### EVOLUTION OF A TAXONOMY OF HYPHOMYCETES

The Italian mycologist, Saccardo proposed a simple and elegant classification of these fungi using primarily features of the asexual propagules (conidia), features of morphology, including colour. The discovery and elucidation of polymorphism in fungi by the Tulasne brothers and its later experimental demonstration and confirmation from studies with pure cultures by de Bary and by Brefeld showed that many moulds are but parts of the 'whole fungus', not the 'whole fungus'. Not merely microscopic study, but experimental proof, is needed to ascertain the precise taxonomy (and nomenclature) of a given fungus. When one considers the innumerable species of fungi described from many parts of the world, from all manner of substrates and diverse habitats, and the many new and interesting ones that are continually being reported and described especially from the tropics, the task of obtaining connections and correlations between the 'states' of the whole fungus may well seem hopeless or impossible. And yet, significant progress has been and is being made in this area too.

A new and bold approach to taxonomy of conidial fungi was initiated by the French mycologist, Costantin, who proposed a classification based on the pattern of development of conidia, i.e. conidiogenesis. The key to this classification is developmental morphology. If the source of Saccardo's data for his classification were bibliographic, those of Costantin came from his own original observations on a limited number of moulds he studied. Costantin's ideas have been tested, modified or amplified by many later students, notably by Vuillemin, Mason, Hughes, Tubaki, Subramanian and others. The careful

and systematic study of Hyphomycetes by Stan Hughes which culminated in his 1953 paper on conidiophores, conidia and classification, and his most valuable study (1958) of numerous types of classical genera form a landmark in the development of fungal taxonomy. Thirty years of further studies on Hughesian lines served to strengthen all that is best in the Hughesian system, and to highlight some of its deficiencies. There were two Kananaskis Conferences on Fungi imperfecti which Bryce Kendrick organised with initiative and imagination, one in 1969 and the other in 1977. The first of these examined conidiogenesis and evolved a synthesis and terminology based on consensus; it also provided definitions of many terms, Kananaskis II highlighted the need to look at conidial fungi (anamorphs) as parts of the 'whole fungus' (holomorph). Both Conferences were a fruitful meeting of minds, an unforgettable experience. The new information on ultrastructure of conidiogenesis which became available following Kananaskis I has, however, led to scepticism about its recommendations. In consequence, once again conidiogenesis is being examined from newer angles (see Subramanian 1983, Descals 1985, Minter 1985), but now with a new awareness of the need to evolve hypotheses concerning relationships with predictive value. We have, then, many variations of the theme. Basically, Vuillemin's ideas have been substantiated in some measure. What is noteworthy is that these were originally derived from understanding of conidiogenesis using the ordinary light microscope, and now have been confirmed by studies using transmission and scanning electron microscopy, fluorescent antibody staining and other techniques. All this has led to refinements in the taxonomy of these beautiful organisms. The proper and precise delineation of similarities and differences is of the essence of taxonomy and these pertain not merely to morphological characters, but also to paramorphological parameters. For this reason, fungal taxonomy is a dynamic area of research in biology. The mycologist uses many new tools and techniques in his attempt to ascertain relationships based on form and metabolism.

Alongside the progress in understanding conidiogenesis and its use as a major parameter in the taxonomy of conidial fungi (anamorphs), there has been significant progress in the taxonomy of fungi that exhibit sexuality and produce what are called perfect states (teleomorphs). The gradual evolution of concepts concerning the nature of ascomata based on developmental morphology (Petrak, Miller, Nannfeldt, Luttrell) coupled

with the elucidation of the apical apparatus and other features of the ascus (Chandefaud) led to refinements in the taxonomy of the Ascomycotina.

The establishment of connections between states of polymorphic fungi, between anamorphs, synanamorphs and teleomorphs, initiated by the brothers Tulasne, to which I referred earlier, continues and forms the basis of our efforts, from another angle, towards the same objective. From a study of the present data on anamorph-teleomorph connections, several anamorph-teleomorph correlations have been suggested. The validity and usefulness of these correlations would depend on their predictive value. The aim is to be able to predict the appropriate niche for anamorphs in a classification of teleomorphs. Some suggested correlations seem to satisfy this requirement, but others need confirmation. I shall cite one such correlation which I believe has definite predictive value, to indicate to you how exciting and fascinating research in this area can be

Among the common moulds, *Penicillium*, *Aspergillus*, *Cephalosporium* (*Acremonium*, in current nomenclature) and *Trichoderma* are four of the very common genera. All of them are best known in their (imperfect) conidial state, though some of them reproduce sexually and have a teleomorph. The teleomorphs belong to the Ascomycotina. The conidia of species of all four genera are produced from conidiogenous cells that are known as 'phialides'. The distinguishing feature of the phialide is that it produces an indefinite number of conidia successively and basipetally from its tip, but it does not itself increase or decrease in length; it may or may not have an evident collarette. Though this is the basic feature of phialoconidiogenesis of species of the four genera, there are differences and I would place *Penicillium* and *Aspergillus* in one group, and *Cephalosporium* and *Trichoderma* in another. The rationale of this grouping is:

(i) *Penicillium* and *Aspergillus*: The conidia are produced in true, persistent chains from within the tip of the phialide. They are dry, not slimy, and are typically air- or wind-borne. Their teleomorphs, where known, belong to the order Eurotiales in the Ascomycotina.

(ii) *Cephalosporium* and *Trichoderma*: The conidia are produced singly (not in true or persistent chains) and in succession from within the tip of the phialide: they remain in slime, forming masses or a loose linear series. They are typically slimy and water- or insect-borne, rarely airborne and that too only when the slime has dried. Their teleomorphs, where known, belong to the order Hypocreales of the Ascomycotina.

I had drawn attention to this distinction nearly twenty years ago, primarily on the basis of observations with the ordinary light microscope. In later years support for this came from work on ultrastructure of conidiogenesis. The internal structure of the phialide, and the essentials of cell wall participation in conidiogenesis are distinct in these two groups of phialidic fungi as seen from studies of ultrastructure of conidiogenesis. Thus, in the *Penicillium* type, the conidia are, in the terminology I have devised, novi-cum penititunicogenous, synechidic and porrectic, whereas in the *Cephalosporium* type, they are porrectic, pseudosynechidic, percurrent, first conidium tunicogenous and later conidia demiseptatunicogenous.

The Eurotiales and the Hypocreales are distinguishable on a number of teleomorphic characteristics, apart from the differing pattern of phialoconidiogenesis.

All this is part of the concept of the heterogeneity of the phialide and of phialoconidiogenesis which I proposed some years ago. This is currently on test. The occurrence of totitunicogenous basipetal conidial chains, and of penititunicogenous basipetal conidial chains, on separate phialides in a monoconidial culture as in *Sagrahamala striatispora* and of the *Penicillium* type and the *Cephalosporium* type of conidiogenesis on separate phialides in the same culture in *Gliomastix murorum* suggest that our knowledge of the phialide is far from complete and the future holds surprises for us. The excellent analysis of "patterns of development in conidial fungi" by Cole and Samson (1979) has added a new dimension to our knowledge of conidiogenesis. More recently, Minter et al. (1982, 1983a, b) have contributed substantially to the discussion of this fascinating subject. The evolution of taxonomy is an endless continuum of corrections and close approximation.

I have believed for some time, intuitively, that the distinction between the *Penicillium* and *Cephalosporium* types of phialoconidiogenesis has taxonomic value and so can be used in taxonomy. There are some who do not think so, although they recognise these two genera, and others on this basis. Their scepticism is, nevertheless, a challenge and must be met: not by evidence from ultrastructure of conidiogenesis which is already available, but by other evidence aimed at clinching the issue. Assaults on ideas must be welcomed as they eventually must lead to further refinements in the ideas, as long as they are justifiable. It is my belief that developmental

morphology, which is the basis of the distinction between the two kinds of phialocondiogenesis, must have its roots in metabolism. Distinctions in metabolism are best and most easily discernible in secondary metabolism. Naturally, I have been looking for pertinent data on secondary metabolism of moulds which are characterised by one or the other of these two modes of phialocondiogenesis. And it is most fortunate that I did find some data fulfilling my expectations.

Although I am not a chemist myself, secondary metabolism is an area of special interest to me and I often wonder why in our country which can justly be proud of having built a tradition of excellence in the field of organic chemistry, with several active schools of research, study of secondary metabolism and secondary metabolites of moulds is not given sufficient attention. At the London School of Hygiene and Tropical Medicine, Raistrick did pioneering and valuable work in this area and reported on innumerable secondary metabolites from moulds. Griseofulvin, which is currently one of the few antifungal antibiotics, is a notable example. The discovery of penicillin, another example of a secondary metabolite revolutionised disease control by antimicrobial therapy. Equally important, it gave an impetus to research on secondary metabolites of fungi. The discovery of gibberellins opened up new vistas of research on growth and reproduction of plants. Many more examples can be cited and are doubtless known to you. However, there is relatively little work on biosynthesis of secondary metabolites.

The most thoroughly worked out sequences in biosynthesis, of particular interest to me, concern the biosynthetic pathways that lead to the production of the  $\beta$ -lactam antibiotics, Penicillin G (produced by *Penicillium*) and Cephalosporin C (produced by *Cephalosporium*). Having noted the specificity in the production of penicillin and cephalosporin, and the distinct taxonomic niches *Penicillium* and *Cephalosporium* occupy in current systems of classification, and the distinct positions their teleomorphs occupy in the Ascomycotina, let me say that all this is perfectly in line with the differences in biosynthetic steps species of the two genera exhibit following synthesis of isopenicillin—as far as our present knowledge goes. Though isopenicillin is produced by both species (*Penicillin chrysogenum* and *Cephalosporium acremonium*) and the steps leading to isopenicillin synthesis are the same, the chemical similarity seems to end there. The biosynthesis of Penicillin N (Cephalosporin N) and thereafter Cephalosporin C occur only in the case of *Cephalosporium*; on the other hand, synthesis of Penicillin G from isopenicillin is typical of

*Penicillium*. I have discussed these in detail elsewhere. I wish I could cite more examples of correlation of taxonomy with secondary metabolism. Innumerable fungal metabolites have been described and even characterised, but there is little work on biosynthetic pathways that lead to production of specific metabolites by different fungal species. For example, trichothecenes are produced by several species of *Fusarium* and are well known mycotoxins. They were, however, first discovered in *Trichothecium roseum*, a fungus in which conidiogenesis is apparently distinct from the phialidic mode that characterises *Fusarium*, *Cephalosporium* and *Trichoderma*. Nevertheless, I believe conidiogenesis in *Trichothecium roseum* must bear some relationship to conidiogenesis in *Cephalosporium*, *Trichoderma* and *Fusarium*, though I do not know how, since *Trichothecium* has a (*Hypomyces*) teleomorph state which belongs to the Hypocreales, like the teleomorphs of *Cephalosporium*, *Trichoderma* and *Fusarium* which also belong to the Hypocreales. Several metabolites closely related to trichothecenes are produced by a range of phialoconidial fungi including *Penicillium*, but correlation of secondary metabolism with conidiogenesis such as the one I have given for *Penicillium* and *Cephalosporium* would be possible only if the biosynthetic pathways are known. There is also often the question of the identity of the fungal species reported on, as misidentifications are not uncommon. No other examples of correlation are known to me. Who cares to investigate biosynthetic pathways and secondary metabolism of moulds? And, what for? The example I have given is an exception and stems primarily from our interest in using knowledge of secondary metabolism of the two mould species (*Penicillium chrysogenum* and *Cephalosporium acremonium*) for production of antibiotics for alleviation of human suffering, for human welfare. All this work is truly remarkable, beautiful, and is the result of admirable teamwork by distinguished biochemists, biophysicists and mycologists. Incidentally, and importantly, all this satisfies the human urge to find out and to know, and the human aesthetic sense of wonder at the apparent simplicity and beauty of form and metabolism of these lowly organisms. These organisms are so ubiquitous that no one might show any concern for them. Thus, when Westling first described *Penicillium notatum*, the mould that settled on Fleming's Petri plate in his laboratory in the hospital at Paddington, no one bothered about his description or, for that matter, that of *Penicillium chrysogenum* by Thom. The importance of these two mould species became obvious later: this, of course, is of the nature of science.

The tropics, the forests, soils and seas, our many ecosystems, harbour countless fungi, many of which remain undescribed and unknown: their potential in biosynthesis of useful metabolites and transformations, such as of steroids, is unimaginable. It is fine to run behind the fashionables in biotechnology and genetic engineering which admittedly have much to offer, but let us not belittle the fruits of the earth, the countless organisms naturally available to us and of potential use straightaway. Our forests are a gold mine of fungal resources. It would not suffice if mycologists and others locate, describe and isolate fungi; they have a responsibility to deposit the cultures in a national collection (alas, there is none!). Our friends, the organic chemists have also a responsibility: to find time to study their secondary metabolism. I need hardly remind them that a great deal of the progress in organic chemistry and organic biosynthetic chemistry is primarily a contribution of mould chemistry and mould secondary metabolism.

### SOIL MYCOLOGY

Let me now say something about my fascination for soil mycology. Many years ago Selman Waksman posed the question; Is there a fungus flora of the soil? And by careful and sustained work he showed that fungi, especially moulds, are an important component of microbiota of soils. That there is a fungus flora of the soil is now well established. Soil harbours a variety of fungi doing many things, interacting between themselves and with other organisms. There are saprophytic fungi breaking down substrates from cellulose to keratin, parasitic species infecting root systems of plants or infecting other organisms such as amoebae and nematodes, fungi participating in mycorrhizal symbioses, and many others. There are fungus-host interactions, fungus-substrate relationships. There is antibiosis, there are synergistic effects. These have been the subject of innumerable studies and we do know a lot about them and yet we do not have answers to many questions. The truth eludes us. In the case of soil fungi, such as species of *Fusarium* and *Verticillium* that cause wilt diseases in plants, it is interesting but intriguing why many of them cause little damage at the infection sites in root systems but proceed to occupy the vasculature in characteristic fashion. This is all the more surprising as these pathogens have the enzyme systems required to break down plant cell walls. Gäumann described them as examples of systemic infection with localized symptoms. In many ways these fungi are unique in their behaviour. And there are many subtle variations within this group,



depending on the particular host species-pathogen *forma speciales* combination. It is as if everything is perfectly in place.

The interaction between pathogen and host and the *modus operandi* of the pathogen in causing wilt have been the subject of continued study. A role for toxins such as lycomarasmine and fusaric acid from *formae speciales* of *Fusarium oxysporum* has been postulated. Fusaric acid, but not lycomarasmine, has been detected *in vivo*. Also, neither shows the host specificity that is typical of the pathogen. A role for cell wall degrading enzymes—cellulases, pectic enzymes—has also been suggested. Neither toxins nor cell wall degrading enzymes explain the disease syndrome fully or satisfactorily. Admittedly, we do not yet know the truth, but the work so far has helped in revealing the inadequacy of our current concepts and in eliminating errors. Host specificity which is typical of these pathogens defies elucidation and explanation. The basis for control of these diseases is still largely disease resistance, but it is not easy to pinpoint the basis of resistance. These fundamental questions remain. They must evoke our sense of wonder as much as they vex us by defiance at our efforts to understand them.

Another interesting fusarioid disease syndrome is the "bakanae" or foolish seedling disease of rice plants. The unusual and conspicuous elongation of internodes, which is a striking symptom of infected rice plants, attracted attention early. Enquiry into this led to the discovery that internodal elongation was due to metabolite(s) of the pathogen—the gibberellins which are secondary metabolites of the *Fusarium* (teleomorph genus, *Gibberella*) causing the bakanae disease. Though many fungal species have been screened, the bakanae fungus *Fusarium moniliforme* remains the only fungal source of gibberellin. This species, then, is unique amongst the *Fusaria*.

Garrett's clear and original analysis of the ecological behaviour of soil-borne fungal pathogens and symbionts is still with us, and the variety and beauty of fungus-root relationships that are known is such that a sense of wonder, a passion and persistence to find out the truth, and a tenacity of purpose are needed for progress to come by.

### EXPLORATION OF MICROFUNGI

Finally, I like to say something about my continued exploration of South Indian microfungi which has remained as an obsession with me for many years. The many collection trips to the Nilgiris brought to light new and

interesting hyphomycetes, and the excitement of discovering something new, something beautiful, the experience of beauty of perpetual novelty, remains with me and gives me a sense of ecstasy that is indescribable. In later years, I was joined by my students in these studies and we collected extensively in the Western Ghats, including the Silent Valley. The joy of sharing of excitement of discovery is a unique experience. The rich and diverse mycoflora of the Western Ghats is clearly a reflection of the variety and diversity of the flora and fauna of the area since the fungi themselves are fastidious in their preferences for hosts and substrates. This work, in fact, reinforced the view of conservationists that the Silent Valley should be protected: the hydroelectric project planned to be put up there must be abandoned. Apart from the extraordinary beauty and diversity of these microfungi, beauty and diversity of form, who knows what beauty and diversity in secondary metabolism lie hidden in them? We have one of the grandest, one of the most wonderful and one of the richest of fungal resources here. It will be a tragedy to annihilate this patch of forest. no, this simply should not be. My thinking about the Silent Valley was clear. My feelings were aroused and I felt this area must somehow be preserved. I communicated my feelings together with my data to the then Prime Minister of India, the late Smt. Indira Gandhi. The Prime Minister's response was immediate: she responded by appointing a Committee to consider this important matter. What happened later is well known. The Committee deliberated for over a year, and its recommendations were positive and were accepted: the Silent Valley was declared a 'national preserve'.

Response of the kind I received from the late Prime Minister is rare. And there was everything in her response that suggested to me that she was herself excited about Truth, Beauty and Goodness, about the plant, animal and fungal wealth and resources of the Western Ghats. The excitement of science, the excitement of discovery is its own reward. But there was also transmission of excitement here. What is more, all this led to positive action being taken to preserve the forest.

During the Silent Valley episode I was asked: Is the mycoflora of the Valley so unique that one might say most fungal species found here are limited in distribution to the Silent Valley? For an answer, let me first quote my friends Stan Hughes and Kris Pirozynski:

"The flora of peninsular India, both fossil and living, shares many elements with the floras of Africa and Australia... The *Syzygium/Meliolina*

association provides a further example of this vicarious relationship of India to its former Gondwanalandian neighbours of Africa and Australia. In all three of these now discrete lands, most of the mesophytic floras have been eliminated, modified or fragmented and forced into more humid montane refugia. In all three, the *Syzygium/Meliolina* associations are found in precisely such locations.

"To students of Indian fungi, especially those following Subramanian's footsteps in the Wynad-Nilgiri, to those puzzled by consistent similarities of the mycota of India, inter-tropical Africa and Australo-Papua/New Zealand, we offer a reminder: the mycological road from Ootacamund winds its way to Mysore through Gudalur, Brisbane and Entebbe".

I should add that our knowledge of distribution of fungi is limited; what we know is limited to a few regions where exploration has been carried out. Vast regions on the globe remain unexplored, and opportunity, encouragement and support for such study are hard to come by these days. And yet, there are those who still work in this area impelled by the passion for such research. Yes, the passion for research, to probe into the unknown, with almost religious fervour. In offering their *Carpologia*, the Tulasnes wrote:

"Yet do not say that the authors have determined to spend the rest of their life in idle leisure, bear witness rather that nothing would ever be dearer to them than to spend all that remains of their eyesight, as far as possible to the greater glory of GOD, in contemplating and humbly interpreting those living wonders, albeit of fungi, with which He in His pity deigned to comfort us exiles in this world."

Charles Tulasne provided the illustrations for the *Carpologia* and they are a masterpiece and have never been equalled. Indeed, it is this passion and devotion and love that marks the great works of the pioneers in mycology and natural history.

More about the global distribution of fungi will be known in the future, but what we must note is that the few pockets of tropical forest that still remain, and the fungal resources they carry, may also vanish unless effective steps are taken to protect them from destruction and denudation.

## EPILOGUE

I believe I have said enough to justify the title of my lecture. I must stop here, but let me say this: That the Academy has asked me to deliver this lecture here at Calcutta has given me added pleasure. This is a great centre of culture, art and science. I have many friends here. It is here that Sri Ramakrishna lived and inspired many, notably Swami Vivekananda. It is here that Art and Literature flowered in the towering personality of Rabindranath Tagore. It is here that Jagadish Chandra Bose and Prafulla Chandra Ray pioneered researches in their special fields. It is here that Raman made his great discovery. Again, it is here that Cunningham carried out his pioneering researches on air spora which are a landmark in the development of aerobiology. The work of these great men is impressive and inspiring. But to the young men and women of the future, let me say this: you must make yourself. Let me quote Andre Gide:

"The path that has to be chosen lies through a wholly unexplored country, where each one makes his own discoveries, and—note this—for himself alone; so that the vaguest track in the darkest Africa is more easily distinguishable... Shady groves allure us and the mirage of perennial springs. Or rather, springs will flow where our desire bid them; for the country only comes into existence as our approach gives it form, and the landscape about us gradually falls into shape as we advance; we cannot see as far as the horizon; and even the foreground is nothing but a successive and changeable appearance.

"And so, Nathaniel, you are like the man who should follow as his guide the light he holds in his own hand."

"Let the *importance* lie in your look, not in the thing you look at."

"The wise man is he who constantly wonders afresh."

"It is not enough for me to *read* that the sand on the seashore is soft. My bare feet must feel it. I have no use for knowledge that has not been preceded by a sensation."

"Every perfect action is accompanied by pleasure. That is how you can tell that it was right for you to do it. I don't like people who pride themselves on working painfully. If their work was painful, they had better have done something else. The delight one takes in one's work is the sign of its fittingness, and the sincerity of my pleasure, Nathaniel, is my chief guide."

Thank you.

## References

1. Chadefaud M 1973 Les asques et la systematique des Ascomycètes; *Bull. Soc. mycol. Fr.* **89** 127-170
2. Cole G T and Samson R A 1979 *Patterns of Development in Conidial Fungi* (London: Pitman) 190 pp
3. Costantin J 1888 *Les Mucedinees Simples* (Paris: Librairie Paul Klincksieck) 210 pp
4. Descals E 1985 Conidia as modified hyphae; *Proc Indian Acad Sci (Plant Sci.)* **94** 209-227
5. Elks J (ed.) 1977 *Recent Advances in the Chemistry of  $\beta$ -Lactam Antibiotics* (London: the Chemical Society, Burlington House) 313 pp
6. Fleming I (ed.) 1973 *Selected Organic Syntheses A Guide-book for Organic Chemists* (London: John Wiley & Sons) 227 pp
7. Gams W 1978 Connected and disconnected chains of phialoconidia and *Sagenomella* gen. nov. segregated from *Acremonium*, *Persoonia* **10** 97-122
8. Garrett S D 1956 *Biology of Root-infecting Fungi* (Cambridge: Cambridge University Press) 293 pp
9. Hammill T M 1981 On *Gliomastix murorum* and *G. felina*, *Mycologia* **73** 229-237
10. Hawksworth D L 1985 Problems and prospects in the systematics of the Ascomycotina; *Proc Indian Acad. Sci (Plant Sci.)* **94** 319-339
11. Heitefuss R and Williams P H (eds) 1976 *Physiological Plant Pathology* (Berlin: Springer-Verlag) 890 pp
12. Hughes S J 1953 Conidiophores, conidia and classification; *Can J Bot* **31** 577-659  
 — 1958 Revisiones hyphomycetum aliquot cum appendice de nominibus rejicendis; *Can J Bot* **36** 727-836  
 — and Pirozynski K A 1985 *Meliolina subramanianu* sp. nov.; *Proc. Indian Acad. Sci. (Plant Sci.)* **94** 347-354
14. Kendrick W B (ed.) 1971 *Taxonomy of Fungi Imperfecti*; (Toronto: University of Toronto Press) 309 pp.  
 — (ed.) 1979 *The Whole Fungus*, 2 Vols. (National Museum of Natural Sciences, National Museums, Canada, Ottawa, and Kananaskis Foundation) 793 pp
15. Luttrell E S 1951 Taxonomy of the Pyrenomycetes; *Univ. Mo. Stud.* **24** 1-120  
 — 1955 The ascostromatic Ascomycetes; *Mycologia* **47** 511-532
16. Miller J H 1949 A revision of the classification of the Ascomycetes with special emphasis on the Pyrenomycetes; *Mycologia* **41** 99-127

17. Minter D W 1985 A re-appraisal of the relationships between *Arthrinium* and other hyphomycetes; *Proc Indian Acad Sci (Plant Sci.)* **94** 281-308  
 —, Kirk P M and Sutton B C 1982 Holoblastic phialides; *Trans Br mycol. Soc.* **79** 75-93  
 —, — and — 1983 Thallic phialides; *Trans Br. mycol. Soc* **80** 39-66  
 —, Sutton B C and Brady B L 1983 What are phialides anyway? *Trans. Br mycol Soc* **81** 109-120
18. Nannfeldt J A 1932 Studien ueber die Morphologie und Systematik der nichtlichnisierten inoperculaten Disco myceten; *Nova Acta R. Soc Scient upsal IV* **8** 368
19. Nelson P E, Toussoum T A and Cook R J (eds) 1981 *Fusarium Diseases, Biology and Taxonomy* (Pennsylvania: The Pennsylvania State University Press) 457 pp
20. Saccardo P A 1986 *Sylloge Fungorum omnium hucusque cognitorum* **4** 1-807 (Pavia)
21. Subramanian C V 1971a *Hyphomycetes*. Indian Council of Agricultural Research (New Delhi) 930 pp  
 — 1971b The Phialide, in *Taxonomy of Fungi Imperfecti*, pp 92-119 ed. W B Kendrick, (Toronto: University of Toronto Press)  
 — 1972a Conidium ontogeny; *Curr Sci* **41** 619-624  
 — 1972b The perfect states of *Aspergillus*, *Curr Sci* **41** 755-761  
 — 1972c Conidial chains, their nature and significance in the taxonomy of hyphomycetes, *Curr Sci* **41** 43-49  
 — 1975 Hyphomycetology in retrospect and prospect; *Presidential Address*, Botany Section, 62nd Session, Indian Science Congress, Proceedings, Part II, 53-78  
 — 1979 Phialidic hyphomycetes and their teleomorphs- An analysis; in *The Whole Fungus* pp. 125-151 (Ottawa: National Museum of Canada)  
 — 1980 Anamorphs, teleomorphs and holomorphs; in *Modern Trends in Plant Taxonomy* pp 44-52 ed P K K Nair (New Delhi Vikas Publ. House)  
 — 1983 *Hyphomycetes Taxonomy and Biology* (London: Academic Press) 502 pp  
 — 1986a The progress and status of mycology in India; *Proc Indian Acad. Sci (Plant Sci)* **96** 379-392  
 — 1986b Excursions into Ascomycetology; *Kavaka* **14** 1-8  
 — and Pushkaran M 1975 Conidium ontogeny in some monophialidic Hyphomycetes; *Kavaka* **3** 77-99
22. Thom C 1930 *The Penicillia* (Baltimore: Williams & Wilkins) 644 pp

- and Church M B 1926 *The Aspergilli* (Baltimore: Williams & Wilkins) 272 pp
23. Tulasne L R 1861 *Selecta Fungorum Carpologia* (I English translation by W B Grove) (Oxford: Oxford University Press) 247 pp
24. Turner W B 1971 *Fungal Metabolites* (London: Academic Press) 446 pp
25. Vuillemin P 1910a Matériaux pour une classification rationnelle des Fungi Imperfecti; *C.R. Acad. Sci Paris* **40** 882-884
- 1910b Les Conidiospores; *Bull Soc Sci. Nancy*, Ser. III, **11** 129-172
- 1911 Les Aleurisporés; *Bull Soc. Sci Nancy*, Ser. III, **12** 151-175



**Divya Darshan Pant** (b. 18 October 1919) did D.Phil. (1950) from University of Allahabad. He is INSA Senior Scientist, Head and Senior Professor of Botany and Dean, Science Faculty, University of Allahabad; Executive Councillor, Allahabad and Kumaon universities.

Pant has established himself as a world authority on *Glossopteris* flora, pteridophytes and gymnosperms in general and cycads in particular. He is also an authority on stomatal development and anatomy of vascular plants. Authored new concepts on classification of gymnosperms, fossil spores and pollen, gametophytic nature of *Rhynia gwynnee-vaughani*, phyletic slide of fern annulus and conifer cone evolution. His work is extensively quoted in reviews and textbooks of botany, palaeobotany and palynology. He has authored 'Cycas and Cycadales'.

Pant is Fellow of Indian Academy of Sciences, National Academy of Sciences (India), Linnean Society of London, Indian Botanical Society, and Palaeobotanical Society; Emeritus Member, International Society of Plant Taxonomists; Hony Member, International Organization of Palaeobotany; Hony Life Member, Cycad Society of South Africa, Australia and USA. He is the recipient of Medallion of Birbal Sahni Institute of Palaeobotany (1971); Birbal Sahni Gold Medal (Indian Botanical Society) (1976), JBS Haldane Oration Award and Medal (1988); Birbal Sahni Centenary Gold Medal (1991); Biennial International Gold Medal (Palaeobotanical Society) (1987); T.S. Sadasivan Lecture Award (INSA) (1990).

*Divya Darshan Pant was elected to the fellowship of the Academy in 1972.*

**Address:** Department of Botany, University of Allahabad, Allahabad 211 002; Tel: (0532) 608064 ; Res: 106, Tagore Town, Allahabad 211 002; Tel: (0532) 600864.



# **SOME THOUGHTS ON THE ALGAL ORIGIN OF LAND PLANTS WITH SPECIAL REFERENCE TO THEIR SPORANGIA, SPORES AND ALTERNATION OF GENERATIONS**

**DIVYA DARSHAN PANT**

*The paper deals with the algal origin of land plants in the light of latest researches on their possible algal ancestors. Concepts about the acquisition of land plant characters like a surface cuticle, stomata leading to intercellular air spaces, conducting cells of xylem and phloem, characteristic sex organs like antheridia and archegonia, sporangia and spores and fixed diplo-haplobiontic life cycle are discussed by taking into account the contributions which have been made on the development of such characters by algae and by living and fossil land plants which seem to indicate the manner in which the algae acquired such land plant characters.*

## **INTRODUCTION**

Directly or in connection with other related topics, the algal origin of land plants has been discussed time and again by well known authorities on algae and primitive land plants or by those whose widebased expertise on diverse plants is well known. It is difficult to mention the names of all such persons but we may mention a few randomly selected ones like Fritsch (1908a, 1945b), Church (1919), Lyon (1906), Kidston and Lang (1920), Bower (1908, 1935), Arber (1921), Zimmermann (1930), Corner (1964), Taylor (1982) and Chaloner (1988). Naturally when one is dealing with such a thoroughly discussed topic, it may appear difficult to present entirely new thoughts or conclusions. However, since the edifice of science is built not only on the discovery of new facts but also on the reinterpretation of older ideas, I am venturing to present my thoughts on this much discussed old topic. Indeed in this connection I am reminded of Tagore's immortal lines in "*Gitanjali*", where he says, "And when old words die out on the tongue, new melodies break forth from the heart and where the old tracks are lost new country is revealed with its wonders", since they are universally applicable and as true in the field of science as they are anywhere else.

## ALGAL ANCESTORS

Most botanists believe that the Chlorophyta alone among the diverse groups of algae could have given rise to the land plants. The reasons for this choice lie in these algae and land plants possessing identical photosynthetic pigments (chlorophyll *a* and *b*), storage products (starch) and cellulose in their cell walls. However, opinion remains divided about the forms of Chlorophyta which gave rise to the land plants. Mentioned among the possible ancestral forms are: (1) Charophyceae (2) *Ulva* of Ulvophyceae and (3) some members of the Chaetophorales.

Charophycean forms like *Chara* (figures 1-3) of Characeae have been favoured by some authors because they possess important land plant characteristics like an apical cell, protonemal stage, multicellular sex organs, axial growth habit, parenchymatous organisation and common cytological and biochemical characteristics. Raven (1977) points out that all these features are important in the homoiohydric gas exchange system which allows land plants to regulate their rate of water loss (Stewart & Mattox 1975).

Other authors (Graham 1984) support the ancestral character of charalean genus *Coleochaeta* of Coleochaetaceae and emphasize its having a prostrate thallus made up of radiating filaments that form a one to three cells thick disc (figure 5,6). In some species the disc grows wider by radial and tangential divisions of its peripheral cells and is thus truly parenchymatous. Sexual reproduction is oogamous and the oogonia have a trichogyne. A brief dormancy period follows fertilization after which the zygote, retained in the haploid thallus, becomes partially enclosed by vegetative cells and its wall is coated with sporopollenin. Lately some evidence has been adduced about food material being passed on to the zygote by placental transfer cells surrounding it (figures 62, 63). Morphological comparison between an Early Devonian or Late Silurian thalloid form called *Parka* (figures 7, 8) and *Coleochaeta* is quite close and the presence of such forms, *Protosalvinia* (figure 11) and others at the time when the pioneers of land vegetation had started their invasion of terrestrial habitats lends colour to the ideas about the ancestral character of *Coleochaete*.

Still other botanists prefer *Ulva* (figure 4) of Ulvales to be the progenitor of land plants and they emphasize its occupying the littoral zone, its capability of existing in a wide range of salinity, its

parenchymatous thallus and homologous isomorphic alternation in suggesting its ancestral character.

Authors who regard the Chaetophorales to be ancestral, trace the lineage of land plants from forms like *Frittschiella* (Fritsch 1945a, see figures 9,10). These Chaetophorales grow in wet muddy situations. Among them *Frittschiella* is heterotrichous with a prostrate system having rhizoids and an upward growing green thallus forming the erect system. Cells of both systems undergo divisions in different planes and the parenchymatous growths so formed are comparable with the rhizomatous and aerial axes of the early vascular plants. Some of the Chaetophorales (*Cladophorella*) are also capable of developing a cuticle and Fritsch believes that members of this group with all their potentialities for developing land plant characters became land plants in one leap as soon as they developed a surface cuticle and lignified xylem.

However, Stewart et al. (1973), Stewart and Mattox (1975) and Mattox and Stewart (1984) point out that the cell biology of *Ulva* and Chaetophorales is very different from that of land plants.

### ACQUISITION OF DISTINCTIVE LAND PLANT CHARACTERS

In spite of the consensus about Chlorophyta having given rise to the land plants most of the present day green algae do not show typical land plant characters like (i) a surface cuticle, (ii) stomata leading into a system of intercellular air spaces, (iii) conducting cells of xylem, (iv) a phloem tissue for conduction of elaborated food materials, (v) sex organs called archegonia and antheridia having a single layered jacket of sterile cells where the archegonia have a single axial row of cells whose basal cell alone retains fertility and where the antheridia show numerous rows of exclusively fertile cells, (vi) sporangia having multilayered or single layered multicelled sterile jackets around a number of spore mother cells each of which meiotically gives rise to a tetrad of four spores, (vii) characteristically marked spores coated with sporopollenin and (viii) diplo-haplobiontic life cycle.

In spite of the general lack of the above features in algae, particularly the green ones which are the most likely ancestors of land plants, a number of botanists have attempted to bridge the gap between the algae and the land plants by explaining the differences between the two groups on the basis of a few concepts which are mentioned as follows:

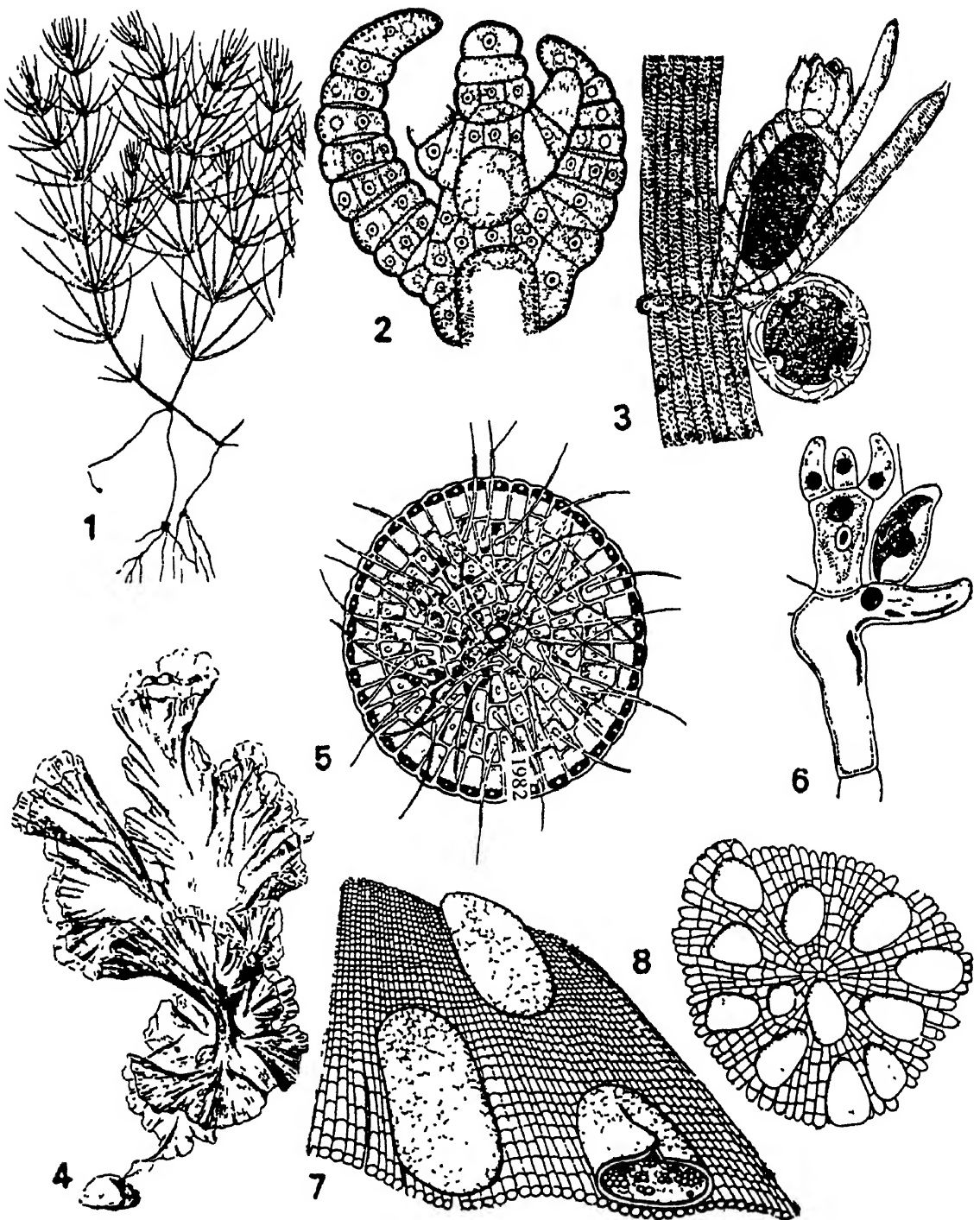


FIG 1-8 *Chara* 1, habit of plant; 2, shoot apex; 3, antheridium and oogonium 4, *Ulva lactuca*, habit of plant. 5,6, *Coleochaete*; 5, habit of plant; 6, oogonium and antheridium. 7,8, *Parkia decipiens*; 7, portion of thallus; 8, reconstruction of thallus. (2-4, after Fritsch, 1935; 7,8, after Taylor, 1982)

### Cuticle

It is believed that the cuticle was developed only when the plant body was exposed to dry air. Fritsch (1935, 1945a, 1945b), Frei and Preston (1961a, 1961b), Hanic and Craigie (1969) have actually described a cuticle in green algae like *Cladophora* and *Chaetomorpha* and also in some Phaeophyceae and Rhodophyceae including *Porphyra*, *Urospora*, *Padina* and *Spongomorpha* (figure 62). The chemical composition of this cuticle has been shown to be different from that of the cuticle of land plants (Hanic & Craigie 1969) but the mere fact that the algae develop a protective layer like the land plant cuticle shows that the algae are capable of developing such resistant layers on their exposed bodies. The chemical composition of this layer could have varied in different phyletic lines and a cuticle with a chemical composition exactly like that of land plants would require the identification of their direct ancestors. Further, in order to say that the algal cuticle is different from the cuticle of early land plants, we would need to determine the chemical composition of the cuticle of diverse early land plant fossils and we would also be required to decide whether the chemical composition of such cuticles has remained unchanged during fossilization.

### Stomata

The stomata of land plants (see figures 19,20) are believed to represent slime pores. A fact which supports this idea is the occurrence of bicelled slime pores in the thalli of *Anthoceros* and *Dendroceros* (Bower 1935, Schuster 1966 see figures 12-18). The prothalli of some ferns like *Ceratopteris* also have similar stomata or slime pores. In this connection it is important to point out that the interconnected air spaces which communicate with the external atmosphere through stomata to form the internal ventilating system of land plants can also be regarded as formed by dried up and transformed mucilage cavities (see Corner 1964, Pant 1965).

### Xylem

The xylem of land plants is believed to have developed from the medullary cells in parenchymatous thalli of the algae which invaded land. Stems of many sea weeds have elongated medullary cells, and xylem-like-scalariform thickenings are actually seen in the thalli of *Sargassum* (Hansteen 1892, see also Fritsch 1945, see figures 21, 22, 23. All these medullary cells may also be serving in the conduction of water and

solutes. However, the walls of scalariform axial cells of *Sargassum* are not lignified although it has been suggested that lignification of their walls to make them waterproof was achieved only after their migration to land with the onset of paucity of water in drier subaerial conditions. All the same the potential of developing a xylem-like tissue seems to have been present in algae even while they were living in water and the derivation of xylem from axial cells of the algae which migrated to terrestrial habitats would thus appear to be quite reasonable.

### *Phloem*

The phloem too is believed to have developed from the medullary cells of parenchymatous algal thalli. A phloem-like tissue showing trumpet hyphae was reported in sporophytes of *Laminaria* as early as 1876 by Reinke. Later work of Oliver (1887), Sykes (1908), Killian (1911), Smith (1939) and others confirmed these observations. Indeed Schmitz and Srivastava (1974) have lately shown that this phloem-like tissue actually contains sieve elements where their longitudinal files form sieve tubes which function in translocation of photoassimilates (see figures 59, 60). Their fine structure too is exactly comparable with that of sieve tubes.

### *Sex Organs*

No structures like the land plant antheridia and archegonia have been reported in any algae although they have been compared with the plurilocular gametangia in the haploid phases of some brown algae (Goebel 1902, Davis 1903, Holferty 1904, Lyon 1904, Schenck 1908, Bower 1935, Fritsch 1945a). However, an important difference marking out such algal gametangia from sex organs of land plants is their being wholly fertile, without any jacket of sterile cells around them whereas land plant antheridia and archegonia are characterised by their single layered jackets of sterile cells which surround their central fertile cells. In this connection it has been suggested that the peripheral cells of homologous land plant gametangia were sterilized to protect the centrally situated fertile cells under land conditions (figure 24A-H). The antheridia, are quite like plurilocular gametangia in having a number of central rows of fertile cells but the archegonia are seemingly different in having only a single axial row of cells inside the jacket and even among these only the basal cell is fertile. However, some abnormal archegonia described in *Mnium* and other archegoniates (see Holferty 1904, Haupt 1926, Bryan 1927 see figures 25-33) have multiple central rows like those of antheridia and

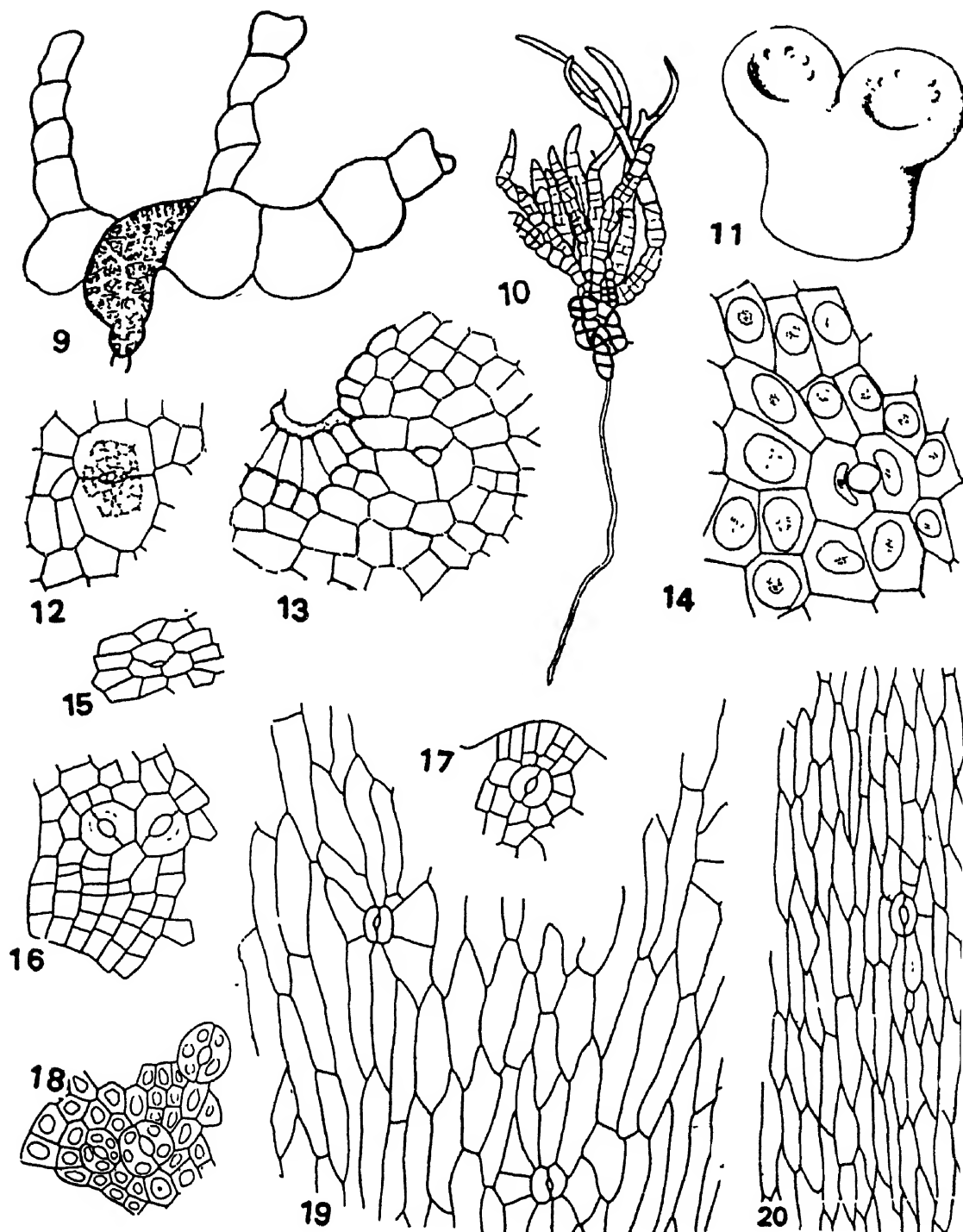


FIG 9-20 9,10, *Früschella* (after Fritsch 1935); 9, prostrate system one cell of which is forming zoospores; 10, habit of a small mature plant, 11, *Protosalvinia*, stomata on ventral surfaces of gametophytic thallus of *Anthoceros*; 12-17, and *Dendroceros*: 18, where guard cells show multiple chloroplasts and ordinary epidermal cells have single chloroplasts; 19, *Lyonophyton rhyniensis*, epidermis from underside of a gametangioophore showing stomata; 20, epidermis of axis of *Rhynia gwynne-vaughanii* 12-14 after Schuster 1984; 15-18, after Bower 1935; 19,20 after Remy & Hass 1991)

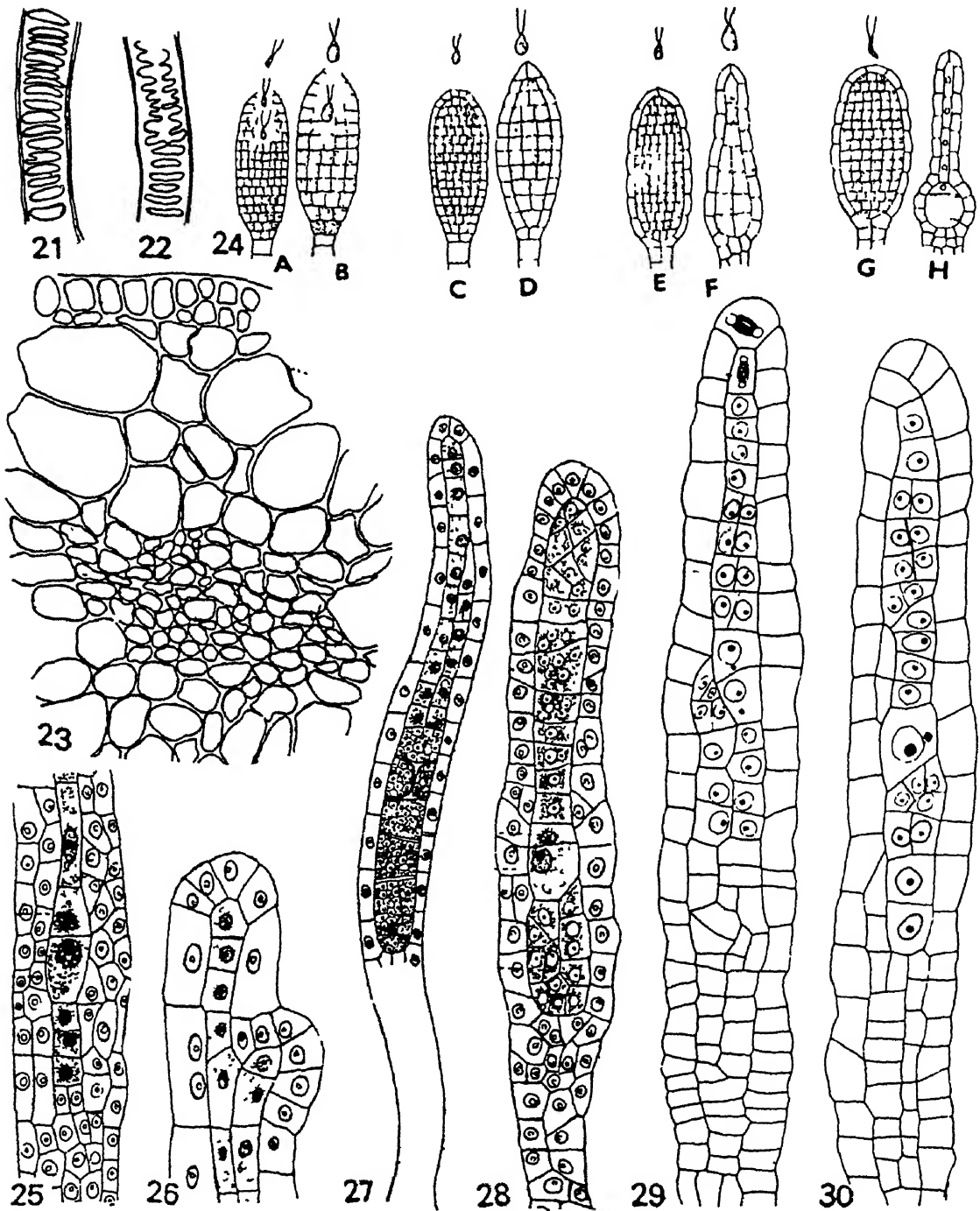


FIG 21-30 *Sargassum bacciferum* (after Hansteen 1892): 21,22, L.S. of medullary cells of thallus showing scalariform thickenings on their longitudinal walls; 23, T.S. of "cladode"; 24, A-H, Diagrams showing hypothetical stages in origin of sex organs of Archegoniatae (based on Davis, 1903), 25-30, *Mnium*, abnormal archegonia (after Bryan 1927)



sometimes more than one enlarged egg-like cell. These abnormal archegonia could be regarded as reversions and they suggest that the archegonia too can be derived like antheridia from plurilocular gametangia. Reduction in the number of fertile cells to a single central row and even in this row the retention of fertility by a single proximal cell in the archegonia was perhaps necessitated by greater requirements of nutrition for the successful development of the progeny in oogamous reproduction.

It has been suggested that the absence of typical plurilocular gametangia among the green algae does not form a serious objection for deriving antheridia and archegonia from such structures since very often even in some green algae of Chaetophoraceae single swarmer producing cells occur in rows and forms like *Draparnaldia* and *Draparnaldiopsis* have entire laterals producing swarmers (Singh 1942) and given the attribute of forming longitudinal partitions in each cell these become comparable as suggested by Fritsch (1945a), with the plurilocular gametangia of brown algae. Fritsch also pointed out that the antheridia of *Chaetonema* are produced by transverse and longitudinal divisions of vegetative cells although unlike the antheridia of land plants each compartment here produces a number of male cells. The derivation of antheridia and archegonia of land plants from the plurilocular gametangia of algae does not, therefore, present any serious problem.

### *Sporangia*

The origin of early land plant sporangia too has been traced from algal sporangia. Out of the two basic types of algal sporangia recognised by Guiry (1978) called meiosporangia and mitosporangia, the land plant sporangia have been compared with the commonest forms of algal meiosporangia, generally called tetrasporangia (figures 40-42). These sporangia too occur likewise in the diploid phase of algae. In making this comparison Schenck (1908) suggested that the spore mother cells of land plants were like the tetrasporangia of algae since both kinds of structures formed tetrads of spores after meiosis (figures 38-39). Kidston and Lang (1921) also emphasized the comparison, which as Fritsch (1945a) points out, "is particularly striking with the tetrasporangia of Dictyotales and diplobiontic Floridaceae".

#### *Differences between algal tetrasporangia and land plant sporangia:*

However, a typical tetrasporangium is fully exposed and produces only a single tetrad of tetraspores whereas the early land plant sporangia like

those of *Rhynia* or *Horneophyton* give rise to a large number of tetrads of spores inside multilayered jackets of sterile cells. No doubt these land plant sporangia were compared by Kidston and Lang (1920, 1921) and Arber (1921) with sori of tetrasporangia found embedded in jackets of peripheral sterile cells in certain pseudoparenchymatous sea weeds (figures 43, 44), although, as Fritsch also points out, the embedded sporangia of parenchymatous sea weeds are never completely surrounded by sterile cells. In spite of his pointing out this lack of exactly comparable structures among algae, Fritsch has himself supported the view held by Kidston and Lang (1920) and Arber (1921), in a slightly modified form, when he suggests that the sori of early land plant sporangia may have been superficial and borne towards the ends of aerial branches. Fritsch (1945a) believes that the change that resulted in the internal layers of parenchymatous thalli (like those of the present day Ectocarpales, e.g., *Punctaria*, *Dictyosiphon*, *Coilodesme*, being given to production of spores, must have been a momentous one although he admits that no adequate explanation can be provided as to its mode of occurrence. The problem of the origin of land plant sporangia has thus been left out without making any cogent suggestions about the manner in which a large number of tetrasporangia of land plants became crowded and at the same time embedded in a peripheral sterile tissue. Another problem which has remained unsolved is the evolution and diversification of land plant spores from those of algae.

The only causal factor which Fritsch (1945a) mentions for the grouping and embedding of sporogenous tissue in early land plants is the secretion of a surface cuticle. It is, however, difficult to understand how the provision of a waterproof protective coat on the exposed surface of the plants which started the invasion of land could have been responsible for the crowding of tetrasporangia and for sinking them deeper in their parenchymatous thalli. Instead the secretion of a surface cuticle could have actually helped such pioneers in retaining their sporogenous cells in their original scatter and superficial position.

*Crowding of sporogenous cells:* As far as the crowding of tetrasporangia is concerned this is already achieved by some brown and red algae. Likewise the Chlorophyta which gave rise to land plants may also have developed similar sori of tetrasporangia.

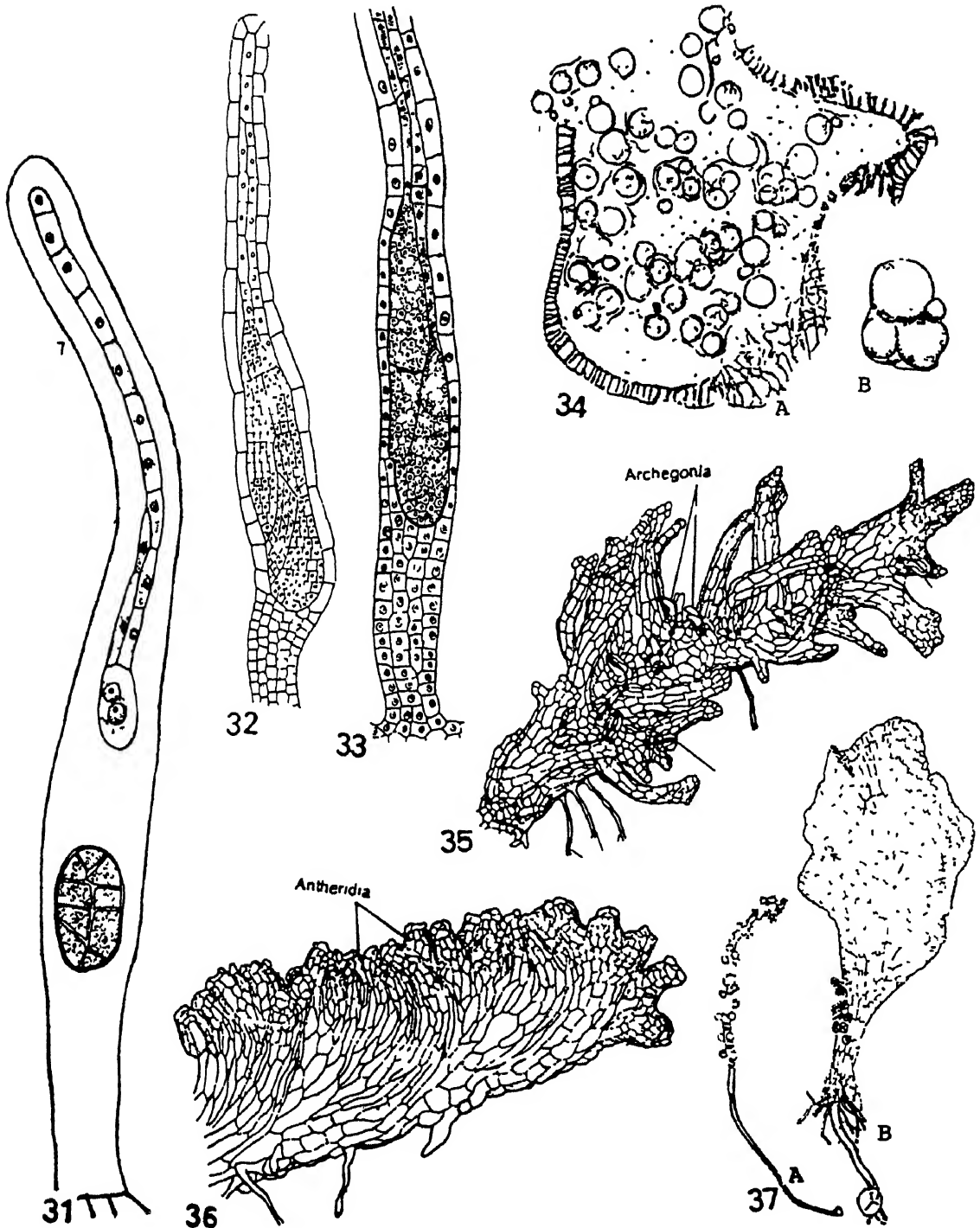


FIG 31-37' 31-33, *Mnium*, abnormal archegonia (after Bryan 1927); 31, abnormal archegonium with a partially biserrate row of neck canal cells and an antheridium below it; 32,33, abnormal archegonia with multiserrate rows of fertile cells towards their bases; 34, *Calamostachys bunneyana*; A, sporangium showing incipient heterospory; B, a single tetrad from A further magnified to show unequal spores (after Scott 1920); 35,36, *Equisetum arvense* portions of female and male gametophytes, respectively (after Duckett 1970); 37, *Platyzoma microphyllum*; A, B, male and female gametophytes, respectively (after Tryon 1964)

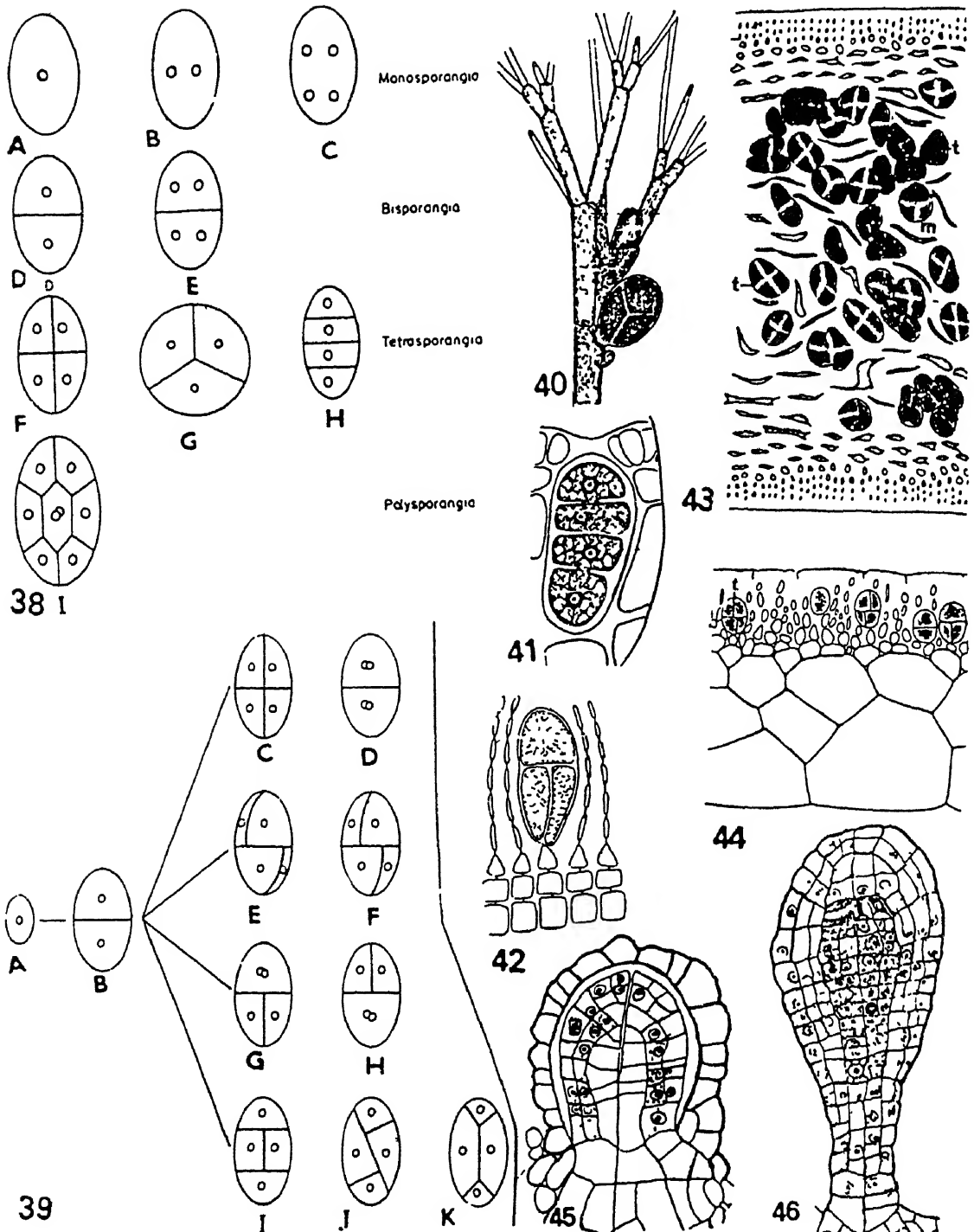


FIG 38-46 38, A-I, Diagrammatic representation of some sporangial types found in Floridiophyceae, 39; A-K diagrammatic representation of the various types of successive cruciate tetrasporangia, 40, *Callithamnion corymbosum*, tetrahedral tetrasporangia; 41, *Cystoclonium purpureum*, zonate tetrasporangium, 42, *Peyssonnelia dubyi*, cruciate tetrasporangium, 43, *Chondrus crispus*, transverse section with mature tetrasporangia; 44, *Chrysumenia ventricosa*-cruciate tetrasporangia; 45, *Notothylas indica*, L.S. of embryo showing commencement of differentiation in amphithecium, 46, *Notothylas levis*, L.S. of embryo showing endothelial origin of sporogenous tissue (38, 39 after Guny 1978; 40-44 after Fritsch 1945; 45, 46 after Pande 1932 and 1934, respectively)

*Embedding of sporogenous cells by a centripetal shift:* On the contrary, it seems reasonable to assume that the drier subaerial conditions on land, which induced the development of a protective surface cuticle or sterile jackets around archegonia and antheridia may also have been responsible for shifting of the sporogenous cells deeper for their protection. This could have been a developmental shift where the superficial layers were diverted to form the protective sterile jackets and the centripetally situated cells became fertile. Such a developmental shift can actually be seen in the sporogonia of modern *Notothylas* whose gametophytic cells with single alga-like chloroplasts could perhaps entitle them to be considered in a discussion of the algal ancestry of land plants. Some species of *Notothylas* like *N. indica* (Pande 1932) produce spores entirely from the amphithecium, other species described by Lang (1907) and Campbell (1908) produce them partly from amphithecium and partly from endothecium (figure 45) and still other species like *N. livieri* (Pande 1934) produce them entirely from endothecium (figure 46). Goebel (1915) and Fritsch (1945a) too had mentioned the centripetal shifting of the sporogenous cells in Anthocerotales but without emphasizing its importance in the evolution of early land plant sporangia. The centripetal shifting during evolution of Anthocerotales seems to be just the reverse of the course of evolution suggested by Bower (1935) for shifting of archesporial cells.

The concept of the developmental centripetal shift of tetrasporangial mother cells in early land plants suggested above could therefore indicate that the elimination of central columns of sterile tissue or columellae from sorogonia or sporangia represents a higher stage of evolution adapted for subaerial existence although Pande (1934) believed that it only showed evolution by reduction. The absence of columellae or central columns of sterile tissue in sporogonia of hepatics or the sporangia of most vascular plants except a few primitive ones like *Horneophyton* and some fossil and trabeculae comparable with bryophytic columellae (see Pant & Srivastava 1962) seemingly confirms the suggestion about the loss of columellae representing a higher stage of evolution.

Pande (1934) also believed that the variations seen in the layers which give rise to the sporogenous tissue in different species of *Notothylas* (either entirely from amphithecium or partly from amphithecium and partly from endothecium or entirely from endothecium) suggested that the genus formed a link between the Anthocerotales and the Hepaticae and he emphasized that the two groups were not fundamentally different. This

may be true but as mentioned above, the genus seems to have a far greater significance in suggesting the manner in which the sporogenous tissue gradually sank into deeper layers of the spore producing parts of the sporophytes. In fact one need not assume it but can actually see it in the sporogonia (sporangia) of different species of *Notothylas*.

*Ontogenetic shifting:* Applying the theory of recapitulation by Haeckel & Von Baer, the ontogeny of primitive land plant sporangia could itself be cited in support of the phylogenetic centripetal shifting of their sorogenous cells like tetrasporangia. To begin with the sporangia of all pteridophytes have superficial initials. Thereafter they embed their sporogenous cells by undergoing periclinal divisions to give rise to a covering layer of primary wall cells and below it the primary sporogenous cells. In eusporangia the primary wall cells divide by periclinal as well as anticlinal walls to produce their many cell thick jackets which deeply embed their sporogenous cells (figures 47, 48). In the higher ferns the phenomenon of reduction results in the primary wall cells of leptosporangia repeatedly dividing only by anticlinal walls to produce single layered jackets of sporangia.

### MEIOSPORES OF ALGAE AND LAND PLANTS

As already mentioned the spores of land plants are closely comparable with the tetraspores of algae. Both kinds of spores are formed from mother cells whose nuclei undergo two divisions of meiosis and the partitions between the spores are formed successively or simultaneously to give rise to tetrads where their arrangement may be linear, cruciate or zonate isobilateral, cruciate or zonate decussate or tetrahedral (figures 38, 39).

In the algal tetrasporangia all the four spores are equal and free from each other. Their surface is typically smooth and altogether lacks any dehiscence marks. Like the surface of algal plant body, the surface of algal tetraspores generally lacks outer sporopollenin coats although such coats have been reported to occur in the zygote walls of Zygnemates, Charales and Coleochaetales (Mattox & Steward 1984). However, there is need of looking for algal tetraspores having resistant sporopollenin coats. After algal tetraspores are set free they germinate to produce haploid plants (gametophytes) and all algae producing tetraspores could therefore be regarded as homosporous.

To begin with, the mother cells of spores in the earliest land plants were likewise giving rise to four equal and free spores. As far as we know

their surface was almost smooth (Chaloner 1970) although it had acquired a resistant coat of sporopollenin obviously for their protection in the drier subaerial environment. The formation of a harder sporopollenin coat may have led to and also necessitated the formation of dehiscence marks. Later on the spores became variously sculptured perhaps to help in their dispersal by various agents available in the changed terrestrial environment, e.g., animals, wind and rain-water. The sculptured coat may have been acquired even for the protection of spores from the agents which dispersed such spores, e.g., the silicious coat of spores of *Isoetes* or the sculptured coats of land plant spores are believed to protect the spores from being digested by animals which seem to be dispersing them (Duthie 1929, Pant & Srivastava 1962, Chaloner 1976, 1984).

*Evolution of Heterospory:* Another change which seems to have come in land plant spores is the formation of unequal spores or heterospory. To begin with the four spores of a tetrad, or the spores in different tetrads inside one and the same sporangium started becoming unequal. The production of such unequal spores by the sporogonia of some mosses, e.g., *Macromitrium* (Fleischer 1920, Ernst-Schwarzenbach 1936, 1938, 1939, 1942, 1944, Ramasay 1979, Vitt 1968) and *Schlotheimia* which give rise to two kinds of gametophytes (Ramasay 1979) is termed anisospory. Pant and Singh (1989) have recently reported similar unequal spores in some Hepaticae and they have suggested that these hepatics too are anisosporous. Similar occurrences of unequal spores are also reported in the sporangia of some fossil pteridophytes like *Calamostachys binneyana* (Scott 1920-see figure 34A, B) *C. americana* (Arnold 1958) *Chaleuria* (Andrews et al. 1974) *Barinophyton*, *Protobarinophyton* (Pettit 1970, Bauer 1980) and some living ones like *Platyzoma* and others (Bower 1926, Tryon 1964-see figure 37A, B). Such production of unequal spores inside one and the same sporangium has been termed incipient heterospory but as Pant and Singh (1989) have pointed out, the two terms anisospory and incipient heterospory have been used by different authors for the same phenomenon.

Among the forms which exhibit anisopory or incipient heterospory (or heterothally) the genus *Equisetum* needs a special mention. As far as the size of its spores is concerned it would appear to be undoubtedly homosporous but according to a few reports some of its spores produce only male gametophytes while others give rise exclusively to female prothalli. However, bisexual prothalli and sex changes are also reported. Accordingly some authors believe that this production of different kinds of

gametophytes seems to depend on the physiological or environmental conditions around the gametophytes. If the spores are grown crowded, the competition of adjacent gametophytes could be depleting the nutrients in the medium, like the paltry internal supply of nutrients in microspores and the resultant gametophytes turn out to be male. On the contrary if the spores are grown relatively far apart they produce female gametophytes but again when the growth of these gametophytes becomes less vigorous due either to their ageing (depletion of their own capacity) or due to the depletion of nutrients around them, they start producing only the antheridia. So far there appears to be no unanimity about the nature of heterothally in the spores of *Equisetum* (see Hauke 1969, Duckett 1970a, 1970b). It is possible that the spores of *Equisetum* represent a stage earlier than anisospory when they attained incipient heterothally but had not yet become typically heterosporous, where specialized megasporangia exclusively produced fewer, larger female spores (megaspores) while others became the microsporangia and produced a much larger number of smaller, male spores (microspores).

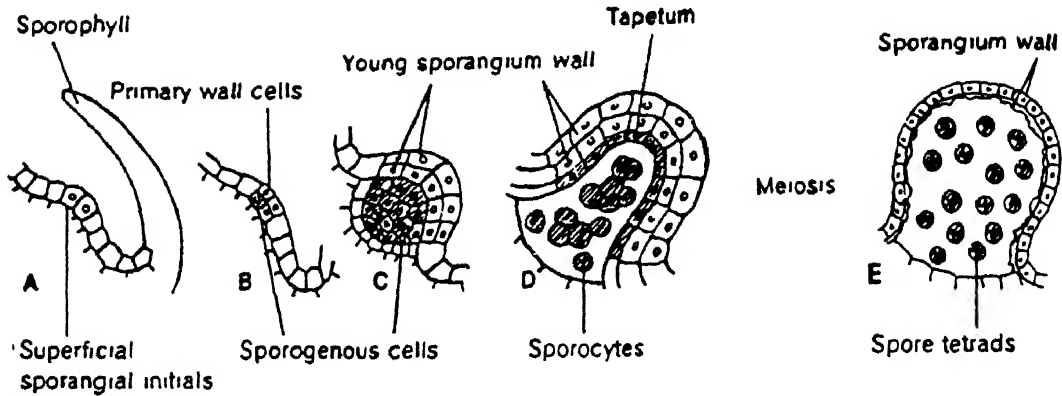
The factors which made some spores larger and left others small are not precisely determined. As mentioned above this factor may have been nutritional as originally suggested by Scott (1920), Shatuck (1910) and Duerden (1929). These authors envisage that the production of larger numbers of spores by a sporangium was cut down by the abortion of some spores so as to lead to the formation of fewer larger spores. However, there are two other alternative explanations for this phenomenon. According to Mogensen (1983) the different sizes of spores in anisoporous bryophytes are genetically controlled while Sussex (1966) believes that a cell "selection" is responsible for the formation of two kinds of spores. Pant and Singh (1989) have, however, suggested that the formation of different sizes of spores could have been controlled by more than one factor.

If we accept the nutritional concept of heterospory, it is natural to assume that larger sized spores (filled with more nutrient material) would give rise to larger gametophytes and after the gametophytes became dioecious the larger spore would be preferred for the production of female gametophytes since their additional store of nutrition could help in the proper development of the next generation. At the same time the smaller spores would be given to the production of male gametophytes which would have no such need for stored food. Indeed the male spores would need to be lighter to give them greater buoyancy.

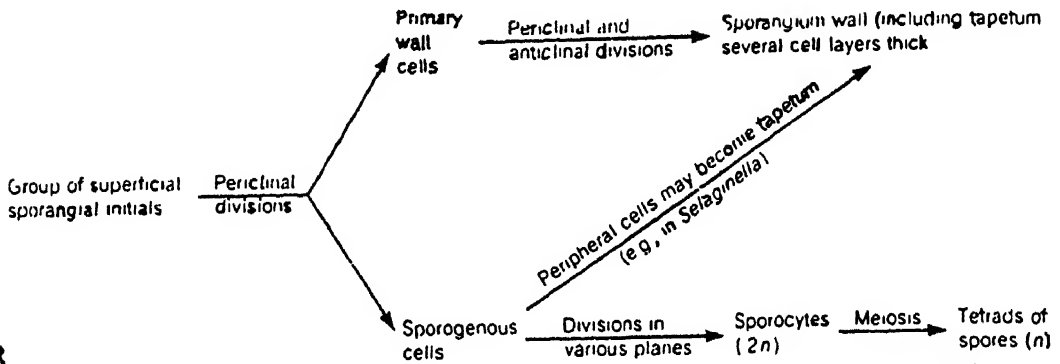


*Obligate tetrads and polyads.* Another change which seems to have come in the spores of plants after their migration to land habitats is their becoming inseparably united in tetrads. Such obligate spore tetrads are among the earliest records of cutinised spores, going back to the Silurian/Ordovician (Gray 1985a, 1985b). Gray attributes these obligate tetrads to bryophytes, since among extant representatives of primitive land plants, which could have existed in Ordovician/Silurian, obligate tetrads occur only in *Sphaerocarpos*, *Cryptothallus* and species belonging to the *Thallocarpus* section of *Riccia*, all of which have dioecious gametophytes. It is possible that the parent plants of the early obligate tetrads were also dioecious and they evolved the strategy of keeping their spores inseparably united in their tetrads in order to overcome the possibility of the isolation of their male and female thalli formed by their dioecious spores in the event of their being widely scattered. Some authors like Chaloner (1988) have pointed out that the attribution of such Late Ordovician to Early Silurian obligate tetrads to bryophytes would imply that dioecious land plants were more primitive than monoecious forms whereas homosporous in early vascular plants would indicate the reverse. However, we need not presume that homosporous early vascular plants were necessarily monoecious. Some of them may have been dioecious like modern *Equisetum*. A form like *Ulva* among the algae which shows isomorphic alternation of generations is dioecious. Indeed after arising from algal ancestors of that kind, some of the first land plants too could have been dioecious.

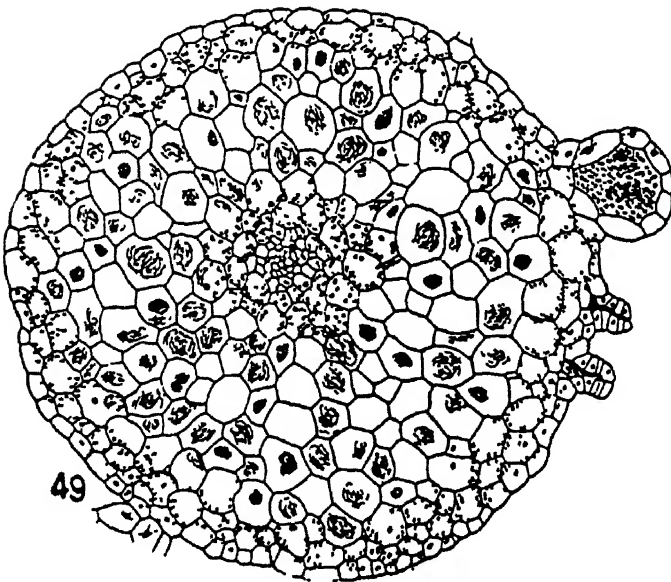
Yet another feature of algal meiospores is their formation in bisporangia as bispores or in monosporangia as monospores. The bisporangia contain two binucleate bispores formed by a single partition between the four nuclei resulting after meiosis while the monosporangia contain a single four nucleate monospore formed after the two meiotic divisions of the nucleus (in this case they remain inside the same cell). Such bispores and monospores have not been described in any bryophyte or pteridophyte but they occur in seed plants particularly in the integumented megasporangia of angiosperms where they are called bisporic or tetrasporic (=monosporic of algae) embryo sacs. However, the bisporic and tetrasporic embryo sacs seem to have been evolved independently in the flowering plants and they are obviously unrelated to algal bispores and monospores.



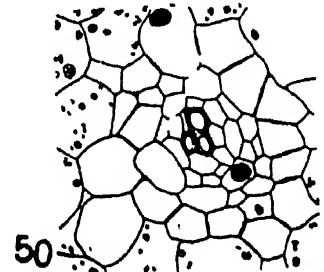
47



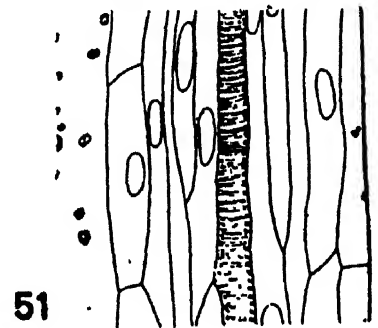
48



49



50



51

FIG 47-51 Structure of a typical vascular plant eusporangium, 48, summary of ontogenetic stages in figure 47 (47-48, after Foster & Gifford 1974); 49-51, *Psilotum triquetrum* (all after Holloway 1939); 49, T.S of a vascularized gametophyte, 50, T.S of stele of a gametophyte further enlarged, 51, L.S of a vascularized gametophyte showing scalariform thickenings in a tracheid.

## DIPLO-HAPLOBIONTIC LIFE CYCLE

While the life cycles of algae can be haplobiontic, diplobiontic or diplo-haplobiontic those of land plants are always diplo-haplobiontic. However, since the land plants are believed to have descended from green algae whose present day representatives have life cycles that are almost always haplobiontic the derivation of their diplohaplobiontic life cycle has posed a serious problem in bridging the gap between the green algae and the land plants.

Some botanists have bridged the gap by suggesting the antithetic or interpolation theory (see Bower 1908, 1935) which assumes that the green algae which gave rise to land plants had a haplobiontic life cycle like that of the present day green algae. The theory postulates that the ancestors of land plants later developed a diploid phase by postponing meiosis and interpolating mitotic divisions to give rise to a diploid phase. Meiosis occurred later in some of the cells of diploid generation and it gave tremendous advantages to the plants as envisaged by Bower.

Other botanists (see Fritsch 1945a) believe that the diplo-haplobiontic life cycle came ready made to the land plants from their diplo-haplobiontic algal ancestors. The supporters of this theory believe that even green algae had evolved a diplo-haplobiontic parallel line of forms like the brown and red algae. This theory derived strong support from the homologous nature of gametophytes in some primitive extant vascular plants like *Psilotum* (figures 49-51), *Tmesipteris*, *Lycopodium*, *Ophioglossaceae* and fossils like *Rhynia gwynne-vaughanii*, which seems to be a gametophyte (see Pant 1962; see figures 52-54). Its sporophyte was presumably a plant of similar form which lacked sex organs, hemispherical projections and adventitious branches but it had terminal sporangia. Another hypothetical model of a possible early land plant is now suggested by Remy and Hass (1991) and they suggest a *Coleochaete*-like ancestor which gave rise to a form having a *Sciadophyton*-like gametophyte and a *Horneophyton*-like sporophyte (see figure 56).

On the contrary Edwards (1980) has asserted that *Rhynia gwynne-vaughanii* is definitely a sporophyte (see figure 58). However, as suggested earlier by Pant (1962), the sporophytic nature of *R. gwynne-vaughanii* continues to remain doubtful since its being a sporophyte can be proved only by showing that its axis which bears the hemispherical projections also bears the sporangia (the two structures should be shown to occur in one organically connected axis). In addition we shall need to

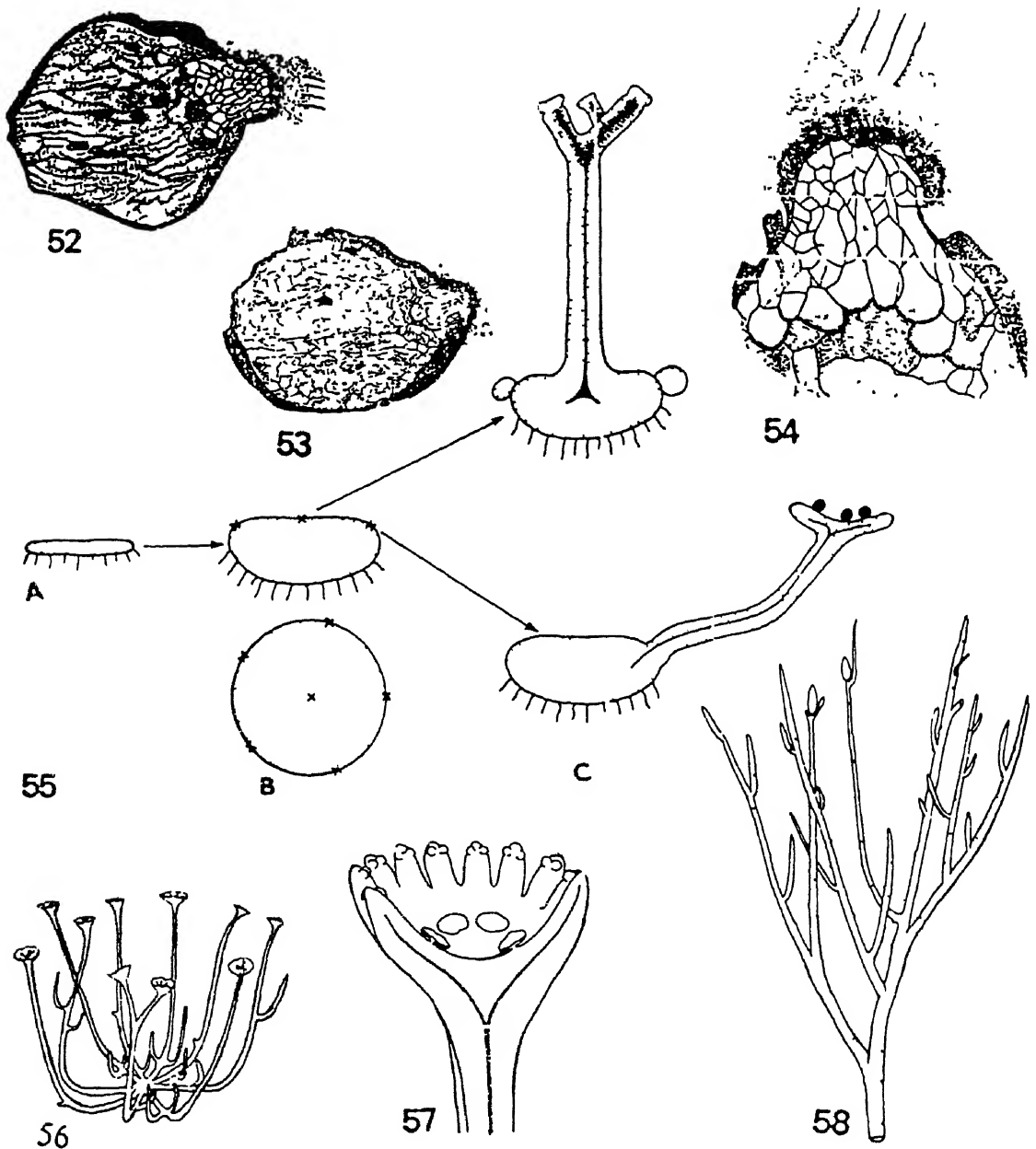


FIG 52-58 *Rhynia gwynne-vaughanu*, 52-53, axes showing hemispherical projects, 54, portion of 52 more magnified to show basal haustorial cells and apical rhizoids (after Pant 1962), 55, A-C, hypothetical scheme of the possible evolution of a land plant archetype, A, initial stage of *Coleochaete* type, B, intermediate stage with tuberos prostrate developments with fertile organs on the flat upper side, C, *Horneophyton*, *Sciadophyton* stages, respectively plants with basal tuberous parts producing aenal axes with terminal fertile parts (after Remy & Hass 1991), 56, *Sciadophyton steinmanni* showing gametophore arising from central disc (from Taylor 1982 after Remy et al 1980); 57, *Lynophyton rhyniensis*, view of split backhalf of a lobed cup of a gametangiophore showing antheridia at tips of lobes and archegonia near base of cup (from Taylor 1982 after Remy 1980), 58, *Rhynia gwynne-vaughanu*, reconstruction of plant (after Edwards 1986)

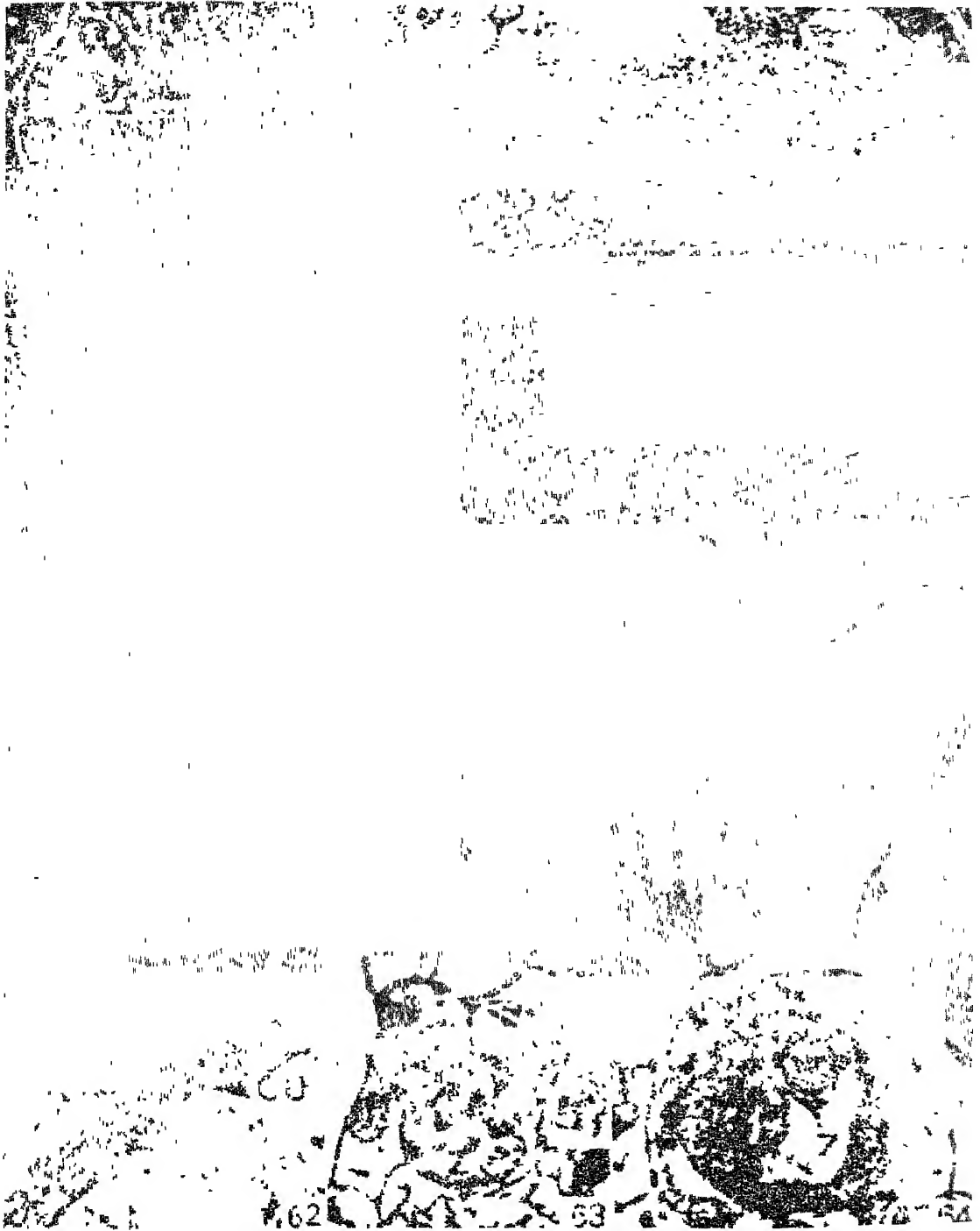


FIG 59-64 *Laminaria groenlandica*, a young sieve element in vivo, showing trumpet shape, vacuoles, plastids and the nucleus, 60, *Laminaria groenlandica*, general view of a mature sieve element near its root wall (59,60 after Schmitz & Srivastava 1974), 61, *Lyonophyton rhyniensis* apical part of median section of a gametangiophore (after Remy & Hass 1991), 62, *Spongomorpha arcta*, cross section of cuticle fragment (after Hanic & Craigie 1969), 63-64, *Coleochaete Pulvinata* zygote with covering cells of jacket (after Graham 1984)

explain the occurrence of archegonium-like structures in axes of *Rhynia gwynne-vaughanii* as shown by Pant (1962) and Le Moigne (1968). We must also explain the nature of the peculiar hemispherical projections which are closely comparable with young sporophytes developing in gametophytes of Psilostaceae (see Pant 1962). Their being regarded as adventitious branches developing under stomata is without any parallel in any other living or fossil vascular plant. As far as I know this has not been done and its being a sporophyte continues to remain doubtful and therefore, as suggested by Pant (1962) and Le Moigne (1968) its being a gametophyte continues to remain a strong possibility.

### CONCLUSION

In the above article I have merely tried to present some ideas which have struck me on the acquisition of land plant characters by the algae in their conquest of land. In doing so even if I have succeeded merely in focussing your attention on this age old problem I feel that I have been amply rewarded but, if in addition, I have been able to convince some of you on the reasonableness of even a few of my thoughts, I feel that I have achieved more than what I had expected.

### ACKNOWLEDGEMENT

I am thankful to Dr D K Chauhan for help with illustrations.

### References

1. Arber E A N 1921 *Devonian Floras, Study of the Origin of Cormophyta*; Cambridge
2. Andrews H N, Gensel P C and Forbes W H 1974 An apparently neterosporous plant from the Middle Devonian of New Brunswick; *Palaeontology* 17 387-408
3. Arnold C A 1958 Petrified cones of the genus *Calamostachys* from the Carboniferous of Illinois; *Contrib. Mus. Paleont., Univ. Mich.* 14 149-165
4. Bauer D F 1980 *Barinophyton citrulliforme* (Barinophytales, incertae sedis, Barinophytaceae) from the Upper Devonian of Pennsylvanian; *Amer J Bot.* 67 1186-1206
5. Bower F O 1908 *Origin of a Land Flora*; London  
 -1926 *The Ferns (Filicales)*, 2. *The Eusporangiate and Other Relatively Primitive Ferns*; (Cambridge: Cambridge University Press)  
 -1935 *Primitive Land Plants*, London

6. Bryan G S 1927 Abnormal sex organs; *Bot, Gaz.* **84** 89-101
7. Campbell D H 1908 Studies on some Javanese Anthocerotaceae II; *Ann.Bot.* **22** 91-102
8. Chaloner W G 1967 Spore and land plant evolution; *Rev Palaeobot. Palynol.* **1** 83-93  
 —1976 The evolution of adaptive features in fossil exine; in *The Evolutionary Significance of the Exine* eds I K Ferguson and J Muller 1-14 *Linnean Soc, Ser* (Academic Press)  
 —1984 *The Biology of Spores*, 17th Ann Mtg. A.A.S.P. Program & Abstracts, 3  
 —1988 The early land plants the saga of a great conquest; *Proc XIV Internat Bot Congr* 301-316
9. Church A H 1919 *Thalasssiophyta and the Subaerial Transmigration*; Oxford Bot Mem., III
10. Corner E J H 1964 *The Life of Plants*, London
11. Davis E J H 1903 The origin of archegonium, *Ann Bot* **17** 477-492
12. Duckett J G 1970a Sexual behaviour of the genus *Equisetum*, subgenus *Equisetum*, *Bot Jour Linn Soc.* **63** 327-352  
 —1970b Spore size in the genus *Equisetum*, *New Phytol* **69** 333-346
13. Duerden H 1929 Variations in megaspore number in *Selaginella*, *Ann.Bot* **43** 451-457
14. Duthie A V 1929 The method of spore dispersal of three South African species of *Isoetes*, *Ann Bot* **43** 411-412
15. Edwards D S 1980 Evidence for the sporophytic status of the Lower Devonian plant *Rhynia gwynne-vaughani* Kidston and Lang, *Rev Palaeobot Palynol* **29** 177-188
16. Ernst-Schwarzenbach M-1936 Zur Heterosporie in der Laubmoos Gattung *Macromitrium*, *Verhandlungen der schweizerischen naturforschenden Gesellschaft* **117** 310-311  
 —1938 Dimorphismus der Sporen und Zwergmannchen- Problem in der Laubmoos Gattung *Macromitrium*; *Ann. Bryol.* **11** 46-55  
 —1939 Zur Kenntnis der sexuellen Dimorphismus der Laubmoos. *Archiv der Julius Klaus-Stiftung für Vererbungsforchung, Sozialanthropologie und Rassenhygiene* **14** 362-474  
 —1942 Weitere Mitteilungen über den sexuellen Dimorphismus der tropischen Laubmoos Gattung *Macromitrium*; *Archiv der Julius Klaus- Stiftung für Vererbungsforchung, Sozialanthropologie und Rassenhygiene* **17** 458-461

- 1944 La Sexualite et la Dimorphisme der Spores des Mousses; *Revue Bryologique et Lichenologique* **14** 105-113
17. Fleischer M 1920 Über die Entwicklung der Zwergmannchen aus sexuell differenzierten Sporen bei Laubmoosen; *Berichte der Deutschen Botanischen Gesellschaft* **38** 84-92
  18. Foster A S and Gifford E M 1974 *Comparative Morphology of Vascular Plants*; 2nd ed. San Francisco
  19. Frei E and Preston R D 1961a Cell wall organisation and wall growth in the filaments of green algae *Cladophora* and *Chaetomorpha*, I. The basic structure and its formation; *Proc. Roy. Soc.*, **B154** 70-94  
—and — 1961b Cell wall organisation and wall growth in the filamentous green algae *Cladophora* and *Chaetomorpha*, II, Spiral structure and spiral growth; *Proc. Roy. Soc.* **B155** 55-77
  20. Fritsch F E 1916 The algal ancestry of the higher plants; *New Phytol.* **15** 233-250  
—1935 *The Structure and Reproduction of the Algae*, I. (Cambridge: Univ. Press)  
—1945a Studies in the comparative morphology of Algae. IV. Algae and the archegoniate plants; *Ann. Bot. (N.S.)* **9** 1-29  
—1945b *The Structure and Reproduction of the Algae II* (London, Cambridge; Univ. Press)
  21. Goebel K 1902 Über die Homologie in der Entwicklung männlicher und weiblicher Sexualorgane; *Flora* **90** 259-305  
—1915 *Organographie der Pflanzen*, 2nd ed. ii. 1 (Jena: Bryophyten)  
—1918 *Organographie der Pflanzen* 2nd ed. ii. 2 (Jena: Pteridophyten)
  22. Graham L E 1984 *Coleochaete* and the origin of land plants; *Amer. J. Bot.* **71** 603-608
  23. Gray J 1985a The microfossil record of early land plants, advances in understanding early terrestrialization. 1970-1984; in *Evolution and Environment in Late Silurian and Early Devonian* eds W G Chaloner and J O Lawson *Phil. Trans. R. Soc. London* **B309** 167-195  
—1985b Ordovician-Silurian land plants: The interdependence of ecology and evolution; *Special papers in palaeontology* **32** 281-295
  24. Guiry M D 1978 The importance of sporangia in the classification of the Florideophyceae; in *Modern Approaches to the Taxonomy of Red and Brown Algae* ed. D E G Irvine and J H Price pp 114-144 New York
  25. Hanic L A and Craigie J S 1969 Studies on the algal cuticle; *J. Phycol.* **5** 80-102
  26. Hansteen B 1892 Studien zur Anatomie and Phylogenie de Fucoideen; *Jahrb. wiss. Bot.* **24** 317-362



27. Haupt A W 1926 Morphology of *Preissia quadrata*; *Bot. Gaz.* **82** 30-54
28. Hauke R L 1969 Gametophyte development in Latin American horsetails; *Bull. Torrey Bot. Cl.* **96** 568-577
29. Holferty G M 1904 The archegonium of *Mnium cuspidatum*, *Bot. Gaz.* **37** 106-126
30. Holloway J E 1939 *The gametophyte, embryo, and young rhizome of Psilotum triquetrum* Swartz, *Ann. Bot. N.S.* **3** 313-336
31. Kidston R and Lang W H 1920 On Old Red Sandstone plants showing structure from the Rhynie Chert Bed, Aberdeenshire, Part III. *Asteroxylon mackiei* Kidston and Lang; *Trans. Roy Soc. Edinb.* **52** 643-680  
—and—1921 on Old Red Sandstone plants showing structure from the Rhynie Chert Bed, Aberdeenshire. Part IV, Restorations of the vascular cryptogams and discussion of their bearing on the general morphology of Pteridophyta and the origin of the organisation of land plants; *Trans. Roy Soc. Edinb.* **52** 831-854
32. Killian C 1911 Beiträge zur kenntnis der Laminarien, *Ztschr. Bot.* **3** 433-494
33. Lang W H 1909 A theory of alternation of generations in Archegoniatae based upon the ontogeny; *New Phytol.* **8** 1-12
34. Lemoigne Y 1968 Observation de archegones portes par des axes de type *Rhynia gwynne-vaughanu* Kidston et Lang. Existence de gametophytes vasculaires au Devonien; *Compt. Rend. Hebd. Seances Acad. Sci.* **266** 1655-1657
35. Lyon F 1904 The evolution of sex organs of plants; *Bot. Gaz.* **37** 280-293
36. Mattox K R and Stewart K D 1984 Classification of the green algae. A concept based on comparative cytology in *Systematics of the Green Algae*, eds D E G Irvine and D M John **27** 29-72 (London: Syst. Assoc. Special)
37. Mogensen G S 1983 The spore in *New Manual of Bryology* R M Schuster (Ed.) 325-342 (Japan: Hattori Botanical Laboratory)
38. Oliver F W 1887 On the obliteration of the sieve tubes in Laminariaceae; *Ann. Bot.* **1** 95-117
39. Pande S K 1932 On the morphology of *Notothylas indica*, *J. Indian Bot. Soc.* **11** 169-177  
—1934 On the morphology of *Notothylas levieri*; *Proc. Indian Acad. Sci.* **5** 205-217
40. Pant D D 1962 The gametophyte of the Psilophytales; *Proc. Summer School of Botany, Darjeeling* 276-361  
—1965 On the ontogeny of stomata and other homologous structure; *Plant Science Series* **1** 1-24

- and Singh R 1989 On the possible occurrence of anisospory in some Hepaticae; *Bot. J. Linn. Soc.* **100** 183-195
41. Pettit J M 1970 Heterospory and the origin of the seed habit; *Biol. Rev.* **45** 401-415
  42. Ramsay H P 1979 Anisospory and sexual diamorphism in the Musci; in *Bryophyte systematics* G C S Clarke and J G Duckett **14** 281-316 (London & New York: Systematics Association Special Volume)
  43. Raven A 1977 The evolution of vascular land plants in relation to supracellular transport processes; *Adv. Bot.* **5** 153-219
  44. Reinke 1876 Beitrag Zur Kenntnis der Tange. *Jahrb wiss Bot.* **10** 317-382
  45. Remy W 1980 Die Generationswechsel der archegoniaten Pflanzen im Übergangsfeld von aquatischer zu terrestrischer Lebensweise; *Argumenta Palaeobotanica* **6** 139-155  
— and Hass H 1991 Gametophyten und Sporophyten im Unterdevon-Fakten und Spekulationen; *Argumenta Palaeobotanica* **8** 193-233
  46. Schenck H 1908 Über die Phylogenie der Archegoniaten und der Characeen'; *Engler Bot. Jahrb.* **42** 1
  47. Schmitz K and Srivastava L M 1974 Fine structure and development of sieve tubes in *Laminaria groenlandica* Resenv, *Cytobiologie.* **10** 66-87
  48. Schuster R M 1965 *The Hepaticae and Anthrocerotae of North America East of the Hndredth Meridian*; (New York: Columbia Univ. Press)  
—1984 *New Manual of Bryology*; (Japan: Hattori Botanical Laboratory)
  48. Scott D H 1920 *Studies in Fossil Botany*. 3rd Ed. 2. London
  49. Shattuck C H 1910 The origin of heterospory in *Marsilea*; *Bot Gaz.* **49** 19-40
  50. Singh R N 1942 Reproduction in *Draparnaldiopsis indica* Bharadwaja; *New Phyt.* **41** 262-273
  51. Smith A I 199 The comparative histology of some of the Laminariales; *Amer. J Bot.* **26** 262-585
  52. Stewart K D and Mattox K R 1975 Comparative cytology, evolution and classification of the green algae with some consideration of the origin of other organisms with chlorophyll *a* and *b.*; *Bot. Rev.* **41** 104-135  
—, — and Floyd G L 1973 mitosis, cytokinesis, the distribution of plasmodesmata and other cytological characteristics in the Ulotrichales, Ulvales, and Chaetophorales: phylogenetic and taxonomic consideration; *J. Phycol.* **9** 128-141
  53. Sussex E M 1966 The origin and development of heterospory in vascular plants; in *Trends in plant Morphogenesis*: 140-171 ed E.G. Cutter, New York

54. Sykes M G 1908 Anatomy and histology of *Macrocystus pyrifera* and *Laminaria saccharina*; *Ann. Bot.* **22** 291-325
55. Taylor T N 1982 The origin of land plants palaeobotanical perspective; *Taxon* **31** 155-177
56. Tryon A F 1964 *Platyzoma*, a Queensland fern with incipient heterospory; *Amer. J. Bot.* **51** 939-942  
— — 1988 The origin of land plants: Some answers, more questions; *Taxon* **37** 805-833
57. Vitt D H 1986 Sex determination in mosses; *Mich. Bot.* **7** 195-203
58. Zimmermann W 1930 *Die Phylogenie der Pflanzen*. Jena



**Baij Nath Singh** (b. 31 July 1914; d. 29 December 1985) did D.Sc. (1947) from University of London, UK. He was Senior Scientific Officer, Department of Soil Microbiology, Rothamsted Experimental Station, England; Assistant Director, Deputy Director (Microbiology) and Emeritus Scientist, Central Drug Research Institute, Lucknow.

Singh did phylogenic classification of amoebae, based on their nuclear division patterns. Found that certain amino acids cause encystment of small, free-living amoebae and that cholesterol feeding renders an avirulent *Entamoeba histolytica* into a virulent form.

The method of isolation of eubacteriolytic Myxobacteria has been named 'Singh plate.'

Singh is Fellow of National Academy of Sciences (India); Hony Fellow, Indian Society for Protozoology; Member, Indian Society for Microbiology, Parasitology and Protozoology. He is the recipient of Silver Plaque (Association of Microbiologists of India) (1979); G.P.Chatterjee Memorial Lecture Award (INSA) (1983).

*Baij Nath Singh was elected to the fellowship of the Academy in 1968.*

## FREE-LIVING SOIL AMOEBAE AS HUMAN PATHOGEN

B N SINGH

The discovery made within the last two decades, that small free-living aerobic amoebae belonging to the genera *Acanthamoeba* and *Naegleria* cause fatal human disease affecting the central nervous system, has changed the whole concept of amoebiasis. The subject of amoebiasis can no longer be defined in terms of a single anaerobic amoeba, *Entamoeba histolytica*.

The present review deals with some of the work of others as well as with my own experience in the subject.

### FIELD SOIL AS ABODE FOR FREE-LIVING AEROBIC AMOEBAE

It is not necessary, in dealing with soil protozoa at the present time, to discuss the views held by Ehrenberg, Dujardin, Stein, Butschli and other earlier workers, regarding the distribution of free-living protozoa. The prevalent idea was that these protozoa, in active (trophic) state, could only exist in fresh water or in the sea.

A fertile field soil contains organic matter, mainly of plant origin, which contains starches, fats, organic acids, proteins and amino compounds. They, therefore, supply nitrogenous and carbohydrate needs of bacteria. Besides nitrogen and carbohydrates, various inorganic salts are present in soil solution which are also needed for bacteria. Small free-living amoebae usually feed on bacteria and, thus, they are well provided with food in soil. Soil reaction also plays an important part in the distribution of microorganisms in soil, but Nasir (unpublished data) found that the lowest pH value at which development took place was 3.9 for amoebae, while they were still active at pH 9.5, which was the highest alkalinity tested (cited by Sandon 1927, p. 55).

Interest in soil protozoa began after the publication of Russell and Hutchinson's (1909) theory of the effect of partial sterilization of soil on the production of plant food. This theory attempted to explain 'soil sickness' as being due to excessive numbers of active (trophic) protozoa which by their phagocytic action restricted the bacterial processes going

on in the soil and the remedial effect of partial sterilization by killing these protozoa by sterilising agents. Although this theory has not been fully accepted (Singh & Crump 1953), it aroused great interest in soil protozoa.

D Ward Cutler and his colleagues, under the inspiring leadership of Sir John E Russell, started pioneering work on soil protozoa at Rothamsted Experimental Station, England since 1919. This work laid the foundation of a new branch of science known as soil protozoology. Cutler (1920) developed a culture method for the count of active and cystic protozoa in soil. Cutler et al. (1922), in their quantitative investigation of the bacterial and protozoan population of a Rothamsted field soil at daily intervals for a year, conclusively showed that large numbers of active amoebae were present, and rapid fluctuations in their numbers and in the bacterial numbers took place daily and seasonally; these were not clearly related to weather conditions. The frequency of occurrence of high numbers of active *Naegleria gruberi* (above 100,000 per gram of soil) was significantly related to that of low bacterial numbers (below 30 millions per gram). Cutler et al (1922) thought that this may be due to the feeding of amoebae on bacteria. Why the numbers of active amoebae were low when the numbers of bacteria were high could not be explained. Ciliates were present in very small numbers as cysts in soil.

A detailed account of the historical development of soil protozoology and the possible role that soil protozoa may play in soil are given by Singh (1960, 1963, 1975).

#### ISOLATION AND CULTURE OF SMALL FREE-LIVING AMOEBAE

At the time when the surveys of bacterial and protozoan numbers in Rothamsted field soil were made, the quality of the bacterial food supply was not considered. It was important to discover whether amoebae may be affected by the quality of bacterial food or, conversely, their numbers may be affected by the proportion of edible and inedible bacterial species, as are available in soil. Singh (1941) investigated the feeding of amoebae of different species of bacteria on non-nutrient agar by devising suitable methods. Bacteria used in these experiments differed widely in their characters. About half of them were eaten by amoebae and showed a range of edibility from some that were readily and completely consumed to others that were partially attacked (Singh 1941, 1942, 1945, 1946). The work on selectivity of bacterial food by amoebae revealed that a species of bacteria, such as *Klebsiella pneumoniae* (*Aerobacter aerogenes*) or

*Escherichia coli*, which was readily and completely eaten by a species of amoeba was also very good food for other species of amoebae and amoeboid organisms. This discovery led Singh (1946, 1955) to the use of non-nutrient agar and a species of readily edible bacteria for the isolation and clonal cultivation of amoebae and amoeboid organisms from soil and other substrates and for the enumeration of their numbers from these substrates. The method of isolation and culture of small free-living amoebae developed by Singh has been universally accepted as the procedure of choice for the isolation and culture of pathogenic and non-pathogenic small free-living amoebae from soil and other substrates and from human cerebrospinal fluid (CSF) and human brain tissues postmortem from cases of amoebic meningo-encephalitis (see Singh, 1975 for the literature). Since small free-living amoebae need no nutrient except a suitable edible species of bacteria supplied on non-nutrient agar, pH 6.6-6.8, it is quite unnecessary to add NaCl,  $MgCl_2 \cdot 6H_2O$ ,  $FeSO_4 \cdot 7H_2O$ ,  $Na_2HPO_4$  and  $KH_2PO_4$  to non-nutrient agar, as done by Neff (1957), and NaCl,  $MgSO_4 \cdot 7H_2O$ ,  $CaCl_2 \cdot 2H_2O$ ,  $Na_2 HPO_4$  and  $KH_2PO_4$ , as advocated by Page (1967a).

Among free-living amoebae, strains of human pathogenic *Naegleria* have been found to be inhibited by 0.5% NaCl incorporated in non-nutrient agar (Singh & Das 1970, Carter 1970, Culbertson 1971), though this has been disputed by Cerva (1978) (see Schuster 1979 for the reference) and Griffin (1983). Thus it is safer to use non-nutrient agar without NaCl or antibiotics for the isolation and culture of amoebae.

From the work described above, it is clear that various nutrient media, as used by different workers for the isolation and culture of small free-living amoebae, are not reliable for use from substrates having mixed microbial population (see Singh 1955, 1975, Singh & Dutta 1984). Non-nutrient agar discourages the growth of inedible bacteria and toxogenic micro-organisms coming from these substrates.

#### DISTINGUISHING CHARACTERS OF *ACANTHAMOEBA* AND *NAEGLERIA*

There is still great confusion whether the pathogenic amoeba, not producing flagellate stage, should be placed in the genus *Acanthamoeba* or *Hartmannella*. Thus Martinez (1983) says that Singh (1952) and Singh and Das (1970) do not recognize *Acanthamoeba* as a distinct genus. Page (1967a) recognizes *Acanthamoeba* as distinct from *Hartmannella*. It is,

therefore, of much interest to trace briefly the developmental history of these two genera.

Alexeieff (1912) created the genus *Hartmannella* on the basis of the absence of 'polar masses' during nuclear mitosis, and defined it as follows: "*Pas de corps polaires dans la division nucleaire Tout (ou presque tout) le material chromatique est employé à la constitution de la plaque equatoriale*". Volkonsky (1931) created a subfamily Hartmannellinae in the family Amoebidae to include only those uninucleate amoebae whose resting nucleus contains a single nucleolus. He placed in it three genera; *Hartmannella* (type *H. glebae* Dobell 1914) in which the spindle is barrel-shaped or cylindrical and the cyst wall is smooth or lightly folded; *Glaeseria* (type *G. testudinis* Ivanic, 1926), with the same spindle shape, but with nuclear division occurring in the cyst; and *Acanthamoeba* (type *A. castellanii*, Douglas 1930) for amoebae with conical, pointed-ended spindle and rough cyst wall. Singh (1952), who emended the genus *Hartmannella* Alexeieff (1912) also recognised it on mitotic division pattern, and included in it those uninucleate amoebae whose resting nucleus contains a single Feulgen-negative nucleolus, and pass through a mitotic process in which the nucleolus disappears and a spindle with chromosomes arranged as an equatorial plate is formed. He (1952) agreed with Volkonsky (1931) on the creation of the genus *Glaeseria* on the basis of the nuclear division occurring in the cyst, but pointed out that spindle shape and cyst character were inadequate for separating the genus *Hartmannella* from *Acanthamoeba*, and rejected the latter genus. Singh (1952) also raised the subfamily Hartmannellinae to the status of a family Hartmannellidae, considering that the nuclear division justified this. The family Hartmannellidae was defined as follows: The resting nucleus has either a single Feulgen-negative nucleolus or several Feulgen-negative nucleoli. During mitosis the nucleolus, or nucleoli disappear, and a spindle with chromosomes arranged as an equatorial plate resembling that found in higher animals and plants develops. The nuclear membrane disappears during mitosis. Amoebae may be uni-or multi-nucleate; no temporary flagella have been discovered. *Hartmannella* Alexeieff (1912) emend. Singh (1952) was made the type genus of Hartmannellidae Volkonsky (1931) emend. Singh (1952). Ray and Hayes (1954) and Adam (1964) agreed with Singh. Pussard (1966), while agreeing with Singh (1952) in considering spindle shape as unsatisfactory for intergeneric differentiation, recognised the genus *Acanthamoeba* on the basis of its wrinkled cyst structure, which he judged to be a decisive character.



It may be stressed that Singh (1952) and Singh and Das (1970) in their classification of the order Amoebida Kent, 1880, based on nuclear mitosis in amoebae without test, did not include locomotive form and behaviour of amoebae.

Page (1967a, b, 1968, 1974) has recognized the genus *Acanthamoeba* and has distinguished it from *Hartmannella* on locomotive form and behaviour and on cystic character of amoebae, although amoebae in both the genera have mesomitotic pattern of nuclear division. According to Page amoebae belonging to *Acanthamoeba*, during active locomotion on glass surface, have broad, anterior hyaline lobopodium from which are produced singly or in twos or threes several hyaline projections (acanthopodia). Cysts are polyhedral or thickly biconvex, wall consisting of more or less polygonal or stellate endocyst and more or less rippled ectocyst. Excystment takes place by removal of operculum at the point of contact between endocyst and ectocyst. Page (1968, p. 25) says "Within the genus *Acanthamoeba*, the cysts, though very distinctive as generic character, are somewhat more confusing as a means of distinguishing species". Amoebae having *limax* locomotive form and moving usually by steady flow with cyst wall smooth and rounded, were known, were included by Page in the genus *Hartmannella*. The use of the term steady flow is not justified because Page has himself produced photographic illustration that *H. hibernica* moves in a highly eruptive manner (see also Singh & Hanumaiah 1979).

Singh and Hanumaiah (1979), who have combined the possible phylogenetic classification of amoebae of Singh (1952) and Singh and Das (1970) and the contributions made by pseudopodial school of taxonomy have recognised both the genera *Acanthamoeba* and *Hartmannella* in the family Hartmannellidae only on locomotive form and behaviour of amoebae, and have rejected the cystic character because *A. glebae* (Dobel 1914) and *A. Invadens* (Singh & Hanumaiah 1979) have rounded or spherical cyst wall without pores or opercula (see also Singh & Das 1970, Singh 1981, Singh & Dutta 1984, Misra & Sharma 1980). In the system of classification of Singh and Hanumaiah (1979) locomotive form and behaviour of amoebae have been used at the generic level.

Singh and Hanumaiah (1979) have defined the genera *Hartmannella* and *Acanthamoeba* as follows:

Genus *Hartmannella* Alexeieff, 1912 emend. Singh, 1952 emend.  
Singh & Hanumaiah 1979

The resting nucleus contains a single Feulgen-negative nucleolus. During mitosis the nucleolus disappears and a spindle with chromosomes arranged as an equatorial plate is formed. In active locomotion amoebae assume *limax* form. No temporary flagella are produced. Type species *Hartmannella hyalina* (Dangeard 1900).

Genus *Acanthamoeba* Volkonsky, 1931 emend. Singh & Hanumaiah, 1979.

During mitosis the Feulgen-negative nucleolus disappears and a spindle with chromosomes arranged as an equatorial plate is formed. Amoebae in active locomotion with broad anterior hyaline lobopodium from which are produced singly or in twos or threes several or many, hyaline projections (acanthopodia). Type species *Acanthamoeba glebae* (Dobell 1914)

*A. glebae* has been recognised as the type species according to the International Rules of Zoological Nomenclature.

Sawyer and Griffin (1975) have considered *A. glebae* as having simpler wall.

It may be pointed out that *Hartmannella* and *Acanthamoeba* are also distinct immunologically whether one uses surface or soluble protein antigens (Visvesvara & Balamuth 1975, Singh & Sharma 1983).

I hope in future the pathogenic amoeba should be called *Acanthamoeba* and not *Hartmannella*.

Page (1981) has created a new genus, *Protacanthamoeba*, for an amoeba isolated from Scottish estuary. It produces acanthopodia during active locomotion, but the cyst wall is smooth and circular to oval in outline without any preformed opercula or pores. The nuclear division, though not studied in detail, is mesomitotic. *Protacanthamoeba* is definitely synonym of *Acanthamoeba* unless it was found that it is distinct immunologically from *Acanthamoeba*.

The family Schizopyrenidae Singh, 1952 emend. Singh & Das, 1970 was created for amoebae whose resting nucleus contains a more or less central Feulgen-negative nucleolus or several Feulgen-negative nucleoli, which during mitosis form 'polar masses'. Nuclear membrane persists throughout division. 'Interzonal bodies' may be present. Amoebae may have more than one nucleus, and some genera may produce flagellate stage.

Singh (1952) recognized three genera in the family Schizopyrenidae for the amoebae showing *limax* locomotive form. Type genus *Schizopyrenus* Singh, 1952.

Feulgen-negative nucleolus dividing during mitosis to form 'polar masses'. Temporary flagellate stage is not produced. Type species *Schizopyrenus russalli* Singh, 1952.

Genus *Naegleria* Alexeieff, 1912 emend. Singh, 1952.

'Polar masses' are formed. Feulgen negative 'interzonal bodies' are present during late stage of mitosis. Temporary flagellate stage with two flagella is produced. Type species *Naegleria gruberi* (Schardinger 1899).

Genus *Didascalus* Singh, 1952.

'Polar masses' without 'interzonal bodies' are present during mitosis. Temporary flagellate stage with two flagella is produced. Type species *Didascalus thorntoni* Singh 1952.

Singh (1952) did not recognize the genus *Vahlkampfia* Chatton and Lalung—Bonnaire, 1912 as a valid genus because Vahlkampff's (1905) *limax* amoeba has distinct 'polar masses' and 'interzonal bodies' during nuclear division. Unless an amoeba is discovered that has 'polar masses' and 'interzonal bodies' and does not produce flagellate stage, *Vahlkampfia* should be considered as synonym of *Naegleria* (see also Singh & Das 1970, Singh & Hanumaiah 1979). Chang (1971) has recognized the genera *Schizopyrenus*, *Naegleria* and *Didascalus* in the family Schizopyrenidae and has rejected the genus *Vahlkampfia*.

Singh and Singh (1966) showed that antisera produced against *S. russelli*, *N. gruberi* and *D. thorntoni* gave immobilization reaction in homologous system but not in heterologous system. Thus these genera are also distinct immunologically. Antisera against human pathogenic *Naegleria* (HB-1 strain) did not react with *D. thorntoni*, as judged by immobilization reaction (Singh & Das 1970). Thus *Naegleria* and *Didascalus* are not only distinct on the presence and absence of 'interzonal bodies' but also immunologically. The statement made by Page (1967a, 1974) that the presence or absence of 'interzonal bodies' is not a distinguishing character for *Naegleria* and *Didascalus* cannot be accepted. There is no justification for Page (1976b), Schuster (1979) and John (1982) to call *D. thorntoni* as *N. thorntoni*. Careful investigations on the nuclear division of *Naegleria* spp. have revealed the constant presence of 'interzonal bodies' (Rafalko 1947, Butt Baro & Knorr 1968, Culbertson,

Ensminger & Overton 1968, Fulton & Guerrini 1969, Singh & Das 1970, Chang 1971, 1974, Gordeeva 1973, Das, Willaert & Jadin 1974, Curson & Brown 1976, and others). The failure to find the consistent presence of 'interzonal bodies' by a few techniques used by them.

Immunologically *Acanthamoeba* and *Naegleria* can readily be distinguished both in terms of surface and soluble antigens (see Singh & Sharma 1983).

It is of interest to point out that Page (1974 p. 174) says : "For the benefit of non-specialists it should be emphasized that, whichever nomenclature they wish to employ, Schizopyrenidae and Vahlkampfiidae (as used here) are exact equivalents, as are *Schizopyrenus* and *Vahlkampfiia*. The difference is over the validity of *Vahlkampfia* and the familial name derived from it". It should be noted that Page (1976a) has included only *limax* amoebae dividing by promitosis in the family Vahlkampfiidae. Singh and Hanumaiah (1979), Singh (1981) and Singh and Dutta (1984) have not only included *limax* amoebae but also those amoebae which during active locomotion are oval, oblong, or ellipse and somewhat elongated or nearly fan-shaped in the family Schizopyrenidae, based on promitotic division

Singh and Hanumaiah (1979) and Singh (1981) have clearly shown, based on their own detailed work and those of other workers, that amoebae, whether they are small or large, or possess different nuclear structures or as uni-or multi-nucleate, or have different locomotive forms and behaviour, or with or without flagellate stage, or whether they are parasitic or free-living fall into three groups in the order Amoebida on the basis of their nuclear mitosis, as suggested by Singh (1952) and Singh and Das (1970). The genera of amoebae that can at present be included in the families Schizopyrenidae, Hartmannellidae and Endamoebidae Calkins, 1933 emend. Singh and Das, 1970, are given by Singh and Hanumaiah (1979), Singh (1981) and Singh and Dutta (1984). This system of classification of amoebae should enable workers to identify known and suspected pathogens, and to differentiate them from other amoebae.

### PATHOGENICITY OF FREE-LIVING AMOEBAE

Several reviews dealing with free-living amoebae causing fatal meningo-encephalitis in humans and in lower mammals have been published (Carter 1972, Chang 1971, 1974, Culbertson 1971, 1981, Duma 1972, Griffin 1978, Jadin 1973, John 1982, Martinez 1976, 1980, 1982, 1983,

Martinez & De Jonckheere 1981, Schuster 1979, Singh 1973, 1975, Singh & Dutta 1984, and others).

It was unthought of that soil amoebae might cause disease in humans. Medical interest in these amoebae was aroused when Culbertson and his colleagues since 1958 conclusively showed that *Acanthamoeba* (strain A-1), now known as *A. culbertsoni* (Singh & Das 1970), contaminant of monkey kidney tissue culture cells, caused acute meningo-encephalitis in mice when trophozoites were administered intranasally. They postulated that a similar disease might occur in humans as the result of swimming or bathing in water heavily contaminated with pathogenic *Acanthamoeba*. Culbertson et al. also found that new isolates of *Acanthamoeba* from a variety of sources produced acute meningo-encephalitis and also chronic granulomatous encephalitis in mice after intranasal inoculation (see Culbertson 1971 for the references). A strain (HN-3), which caused chronic granulomatous disease, was identified as *A. rhysodes* (Singh 1952).

Fowler and Carter (1965) were the first to report four fatal human cases of acute pyrogenic meningitis caused by free-living amoebae from Australia. Culture of the brain and meningeal exudate from all cases yielded no bacteria, nor were tubercle bacilli, torula or viruses isolated. Fowler and Carter (1965) thought that these cases were probably due to *Acanthamoeba*. They suggested that the invasion of the brain and the meninges by amoebae was via the nasal mucosa. Butt (1966) and Patras and Andujar (1966) in the USA reported fatal human cases of meningo-encephalitis and thought that they were due to *Acanthamoeba*. Butt (1966) gave the name for the disease as primary amoebic meningo-encephalitis (PAM). All these cases, supposed to have been due to *Acanthamoeba*, are now known to be due to *Naegleria*.

Carter (1968) first isolated in Australia an amoeba-flagellate from two fatal human cases of amoebic meningo-encephalitis. According to Kudo's (1954) system of classification, Carter (1968) recommended its inclusion in the genus *Naegleria* in the family Naegleridae. In the same year Butt et al. (1968) and Culbertson et al. (1968) also reported the isolation of an amoeba-flagellate from a fatal case of Caucasian male. Both the groups of workers placed this amoeba in the genus *Naegleria*, as defined by Singh (1952), and in the family Schizopyrenidae Singh (1952). In a later study of two strains of amoeba-flagellates from fatal human cases, Carter (1970) did not find the presence of 'interzonal bodies' during

mitosis. He (1970) has, in accordance with Page (1967a), now put the amoeba in the family Vahlkampfiidae, genus *Naegleria* Alexeieff, 1912 emend. Calkins, 1913, *N. fowleri* sp. nov. Singh and Das (1970) named the human pathogenic strain (HB-1) of *Naegleria*, obtained from C.G. Culbertson, as *Naegleria aerobia* on the basis of aerobic nature of the organism. Carter (1972, pages 202-3) says—"A direct comparative study of *Naegleria fowleri* and *Naegleria aerobia* is yet to be made, but Singh and I both agree (personal communication, January 1971) that as far as can be judged from their published descriptions they are identical, with one possible exception, viz. the interzonal body described in the mitotic figure of *Naegleria fowleri*. However, if one believes with Page (1967a) that the interzonal body is too capricious a structure to be used as a diagnostic feature, this difference is unimportant and the two species must be considered on present evidence as a single species. In this case *Naegleria aerobia* should become the junior non-valid synonym as the journal in which it was first named (Singh & Das 1970) was published and issued after which I named *Naegleria fowleri* (Carter 1970)".

Singh (1972) pointed out that Carter (1970) described *Didascalus fowleri* because of the absence of 'interzonal bodies' during mitosis. Singh (1973) also pointed out that if the presence of 'interzonal bodies' in Carter's amoeba is found, *D. fowleri* should become the junior non-valid synonym of *N. aerobia*. It has been shown by Chang (1974) beyond doubt that the two strains of *Naegleria* isolated by Carter from human cases have 'interzonal bodies'. There is no such rule in International Rules of Zoological Nomenclature that a wrongly described amoeba should take priority. Chang (1971) suggested the name *N. invadens* for pathogenic strains of *Naegleria* on the presence of 'interzonal bodies'. *N. invadens* is definitely synonym of *N. aerobia* according to International Rules of Zoological Nomenclature.

Since the work of Fowler and Carter (1965), fatal human cases of amoebic meningo-encephalitis have been reported from Australia, Belgium, Czechoslovakia, Great Britain, India, Ireland, Japan, South Korea, New Guinea, New Zealand, Nigeria, North Western Mexico, Panama, Puerto Rico, the USA, Venezuela, Zambia and other places (see Martinez 1983 for the references).

Twenty-three cases of fatal human meningo-encephalitis caused by *Acanthamoeba* have been reported on immunologic studies of human sera or human brain sections after tissue fixation and also on the presence of

cysts (Martinez 1980, 1982). As pointed out before, cystic character is not reliable for identifying *Acanthamoeba*. Several species of *Acanthamoeba* have been implicated in the human disease. Culbertson (1981) doubts the ability of immunological methods in determining the species of *Acanthamoeba* in fixed human brain section. He (1981) has rightly pointed out that retrospective immunologic tests on fixed amoebae in sections from autopsy should not be regarded specific except for the genus and has emphasized the importance of making cultures from patients who may be suffering from meningo-encephalitis caused by *Acanthamoeba*. By the use of gel diffusion presipitin and immunoelectrophoresis tests, it has been found that the valid species of *Acanthamoeba* are related (see Singh & Sharma 1983). Willaert, Stevens and Healy (1978) claim that a species of *Acanthamoeba* can be identified in fixed human brain sections by the use of antisera produced against highly purified plasma membranes of *Acanthamoeba* spp. (Singh & Dutta 1984).

Callicott et al. (1968) isolated *A. astronyxis* from CSF of a patient with meningitis that remitted spontaneously. Martinez (1980) has pointed out that, since CSF was not examined for motile amoebae, the organism may have been contaminants because free-living amoebae were being investigated in that laboratory. Pan and Ghosh (1971) saw motile amoebae in CSF of two patients that recovered. Cleland et al. (1982) have stated that on three occasions *A. rhysodes* was seen in CSF and cultured from a man in Nigeria who had a five year history of excessive sleeping that resulted in a confusional illness with convulsions. He made a partial recovery from the disease. There was rising serum titre of immobilizing antibody against the strain of amoeba isolated. They suggest that it was a case of chronic amoebic meningo-encephalitis. Recently, two cases of fatal human acute amoebic meningo-encephalitis have been reported from Bombay (Gogate et al. 1984). Amoebae were seen in CSF of both the patients. From one case *A. rhysodes* was isolated and from the other, *A. culbertsoni* on non-nutrient agar seeded with *Escherichia coli*. Large numbers of amoebae were seen in brain sections of both the patients, but no cysts could be seen. Both the species of amoebae were highly pathogenic to mice when administered intranasally.

Since the isolation of *N. aeroba* from fatal human amoebic meningo-encephalitis cases from Australia and the USA in 1968, amoebae cultured from human CSF or brain tissue postmortem have turned out to be *N. aerobia*. Strains of *N. aerobia* isolated from fatal human cases have been reported to be similar immunologically, as judged by gel diffusion

and immunoelectrophoresis techniques (see Visvesvara & Healy 1975 and others). According to John (1982), 108 human cases of *N. aerobia* infection have been reported. Martinez (1983) mentions about 130 cases due to *N. aerobia*. A few more human cases have been reported in 1984.

Most of the reported cases of *N. aerobia* are from developed countries rather than from developing countries. This may be due to greater awareness of the disease in the developed countries and may not be due to greater incidence (see John 1982, Martinez 1983).

According to Martinez (1980, 1982, 1983) PAM due to *N. aerobia* is an acute, haemorrhagic necrotising meningo-encephalitis which usually occurs in previously healthy young adults with a recent history of water sport activities. The most affected areas of CNS are base of frontal lobes and cerebellum. The portal of entry of amoebae is through olfactory neuroepithelium. Martinez has suggested the name granulomatous amoebic encephalitis (GAE) for the disease caused by *Acanthamoeba*. It has been reported in chronically ill and debilitated persons and immunologically impaired patients, some receiving immunosuppressive therapy. In some patients immunodeficiency has not been demonstrated. The portal of entry of amoebae to the brain is not properly understood. It appears to be haematogenous probably from a primary focus in the skin (skin ulcers) or the lower respiratory tract (lung) or eyes (corneal ulcers). The clinical picture of *Acanthamoeba* infection in humans is generally chronic. The most affected areas are the midbrain and posterior fossa structures. Out of 23 reported cases, three did not have a typical granulomatous reaction.

The exact mechanisms of pathogenesis of *N. aerobia* is not properly understood. According to some workers amoebae attack cells directly and according to others they produce cytopathic agents (see Chang 1979, John 1982 for the references).

Two cases of fatal human meningo-encephalitis caused by free-living amoebae have been described in which the amoebae in brain sections could not be identified immunologically as *Naegleria* or *Acanthamoeba* (Duma et al. 1978, Martinez et al. 1980). Duma et al. (1978) found in stained brain sections various stages of mitosis in the amoebae. They say (p.469) — "No interzonal bodies characteristic of *Naegleria* were present. Nuclear division was premitotic and the nuclear membrane persisted throughout division. The cysts in sections were rounded and characterized by a prominent, thick, slightly wrinkled



poreless wall often composed of concentric layers partially split." Thus for the characterization of pathogenic amoebae, it is of great importance to isolate them in culture.

Eleven cases of human amoebic keratitis due to *Acanthamoeba* have been reported from the USA, Great Britain, Germany and Holland (see Bos et al. 1981, Ma et al. 1981 for the references). The causative amoeba has been named either *A. Polyphaga* or *A. castellaaii*. According to Culbertson (1981, pages 40-41)— "*Acanthamoeba* and cysts thereof have been demonstrated in human corneas, in and around corneal ulcers, which may proceed to destruction of the globe. Whether the amoebas initiate the ulcer, or invade in secondarily, is a question, I believe. In experimental attempts to cause amoebic corneal ulcers in rabbits, I was unable (unpublished experiments) to produce them using *A. culbertsoni* following corneal epithelial scarification".

#### EPIDEMIOLOGY OF PATHOGENIC FREE-LIVING AMOEBAE

Destruction of *N. aerobia* by normal human serum *in vitro* has been reported by Carter (1970) and Culbertson (1971). Rowan-Kelly et al. (1980) and Holbrook et al. (1980) have shown that absorption of specific antibody from the serum did not remove the amoebicidal activity. Lysis of *N. aerobia* by human serum is due to activation of complement and that the alternative complement pathway can be directly activated by the trophozoites. In humans complement activation via the alternative pathway may inhibit the spread of *N. aerobia* from the CNS to other parts of the body. Ferrante and Rowan-Kelly (1983) have also shown that normal human serum (NHS) contained an amoebicidal property for *A. culbertsoni* (strain A-1). Repeated absorption of NHS with amoebae did not remove the amoebicidal activity, indicating that specific antibody is not required. *A. culbertsoni* (A-1 strain) was able to activate the alternative complement pathway. Similar results were obtained in case of a non-virulent strain of *A. culbertsoni* and *A. rhyodes* (strain HN-3). Thus both *N. aerobia* and pathogenic species of *Acanthamoeba* are unable to spread haematogeneously in humans.

In experimental infections with pathogenic *N. aerobia* and *A. culbertsoni* intranasally in mice, the amoebae multiply rapidly and spread into the cerebrum and cerebellum, causing fatal meningo-encephalitis. Since pathogenic *Acanthamoeba* is much more commonly found in nature as compared to pathogenic *Naegleria*, one would expect that the

trophozoites of the former should enter human nose much more often than the trophozoites of the latter when humans are exposed to polluted waters. It appears that people working on human meningo-encephalitis caused by free-living amoebae do not realize that the trophozoites cannot swim in water. They are bound to round up and sink to the bottom. Unless one inhales mud from the bottom, there is rarely any chance of amoebae entering the nose and causing fatal disease affecting the central nervous system. This seems to be the case in *Acanthamoeba* meningo-encephalitis. It appears that due to a compromised immune system in chronically ill and debilitated persons and immunologically impaired patients, some receiving immunosuppressive thereby, amoebae enter the brain haematogeneously probably from human skin and corneal ulcers. Thus Martinez (1980) considers *Acanthamoeba* as an opportunist organism.

Singh and Das (1972b) showed that intranasal inoculation of mice with flagellate stage of two strains of *N. aerobia*, one from fatal human case in the USA (strain HB-1) and the other isolated by Singh and Das (1972a) from a sewage sludge sample in Lucknow, caused fatal meningo-encephalitis of cent per cent of the mice in 3-6 days when 2000-2500 flagellates were given intranasally per mouse. Even 50 flagellates/mouse caused the death of two out of eight mice. It was clearly pointed out by Singh and Das (1972b) that the flagellates get converted into amoebae in the nose, the latter causing the death of mice. It appears that Culbertson (1981) does not seem to appreciate the importance of this finding in the epidemiology of human meningo-encephalitis caused by *N. aerobia*, although he found no difference in mouse pathogenicity between the two forms. The flagellates have rigid oval shape and swim very actively in water. They are the infective stage because the trophozoites are hardly able to enter human nose from polluted waters. Thus *N. aerobia* cases are much more common compared with *Acanthamoeba*.

A proper investigation on the aetiology of human meningo-encephalitis cases in developing countries is most likely to reveal that the disease caused by *N. aerobia* is not rare. In every case of purulent (suppurative) meningitis, in which bacteria are not found, free-living amoebae should be suspected as aetiological agent. Examination of fresh CSF may show motile amoebae, making this observation mandatory for positive diagnosis and immediate treatment.

## CHEMOTHERAPY OF HUMAN MENINGO-ENCEPHALITIS CAUSED BY FREE-LIVING AMOEBAE

Drugs used to cure *Entamoeba histolytica* infection in man are ineffective in the case of free-living amoebae. There is still complete absence of chemotherapeutic agents for the human disease caused by *Acanthamoeba*. Amphotericin B and miconazole intravenously and intrathecally appear to be the drug of choice for pathogenic *Naegleria* (see Culbertson 1981, Martinez 1983 for the literature).

I should like to suggest that a centre for the study of aetiology of human meningitis cases, especially in children, is most desirable in India. Such studies are likely to reveal that some of the suspected cases due to virus, and not responding to antibiotic treatment, are in fact due to small free-living aerobic soil amoebae.

### References

1. Adam K M G 1964 A comparative study of hartmannellid amoebae; *J Protozool* 11 423-230
2. Alexeieff A 1912 Sur les caractères cytologiques et la systematique des amibes du groupe *limax* (*Naegleria* nov. gen. et *Hartmannia* nov Gen) et des amibes du vertébrés (*Proctamoeba* nov. gen.); *Bull. Soc Zool Fr* 37 149-68
3. Bos H J, Volker-Dieben H J M and Kok-van Alphen C C 1981 A case of *Acanthamoeba keratitis* in the Netherlands; *Trans Roy Soc Trop Med Hyg* 75 86-91
4. Butt C G 1966 Primary amoebic meningo-encephalitis; *New England J. Med.* 274 1476-76
5. —, Baro G and Knorr R W 1968 *Naegleria* (sp.) identified in amoebic encephalitis; *Amer. J. Clin. Pathol.* 50 568-74
6. Calkins G N 1913 Genera and species of Amoeba; *Trans. 15th Int. Congr. Hyg. & Demog.* pp 1-19
7. — 1933 *The Biology of Protozoa* 2nd ed. (London: Baillière, Tindall and Cox.)
8. Callicott J H, Nelson E C Jones M W, Dos Santos J G, Utz J P, Duma R J and Morrison J Y 1968 Meningoencephalitis due to pathogenic free-living amoebae; *J. Amer Med. As.* 206 579-82
9. Carter R F 1968 Primary amoebic meningo-encephalitis: Clinical, pathological and epidemiological features of six fatal cases; *J. Path. Bact.* 96 1-25

- 1970 Description of a *Naegleria* sp. isolated from two cases of primary amoebic meningo-encephalitis, and of the experimental pathological changes induced by it; *J. Path* 100 217-44
- 1972 Primary amoebic meningo-encephalitis. An appraisal of the present knowledge; *Trans. Roy. Soc. Trop. Med. Hyg.* 66 163-213
10. Chang S L 1971 Small free-living amoebas: cultivation, quantitation, identification, classification, pathogenesis and resistance; in *Current Topics in comparative pathology* Vol. 1 (New York & London: Academic Press, Inc.) pp 201-54
  - 1974 Etiological, pathological, epidemiological and diagnostical considerations of primary amoebic meningoencephalitis; *CRT Crit. Rev. Microbiol.* 3 135-59
  - 1979 Pathogenesis of pathogenic *Naegleria amoeba*; *Folia Parasit.* 26 195-200
  11. Chatton E and Lalung-Bonnaire 1912 Amibe *luxax* (*Vahlkampfia* n. gen) dans l'intestin humain; son importance pour l'interpretation des amibes du culture; *Bull. Soc. Path. Exot.* 5 136-143
  12. Cleland P G, Lawande R Y, Onyemelukwe G and White H C 1982 Chronic amoebic meningoencephalitis. *Arch. Neurol.* 39 56-57
  13. Culbertson C G 1971 The pathogenicity of soil amoebas; *Ann Rev Microbiol* 25 231-254
  - , 1981 Amoebic meningoencephalitis; *Antibiotics Chemother.* 30 28-53
  - , Ensminger P W and Overton W M 1968 Pathogenic *Naegleria* sp —Study of a strain isolated from human cerebrospinal fluid; *J. Protozool.* 15 353-363
  14. Curson T M and Brown T J 1976 Identification and classification of the aetiological agents of primary amoebic meningoencephalitis; *N.Z.J. Marine and Freshwater Res.* 10 245-62
  15. Cutler D W 1920 A method of estimating the number of active protozoa in the soil; *J. Agric. Sci.* 10 137-141
  - , Crump L M and Sandon H 1922 A quantitative investigation of the bacterial and protozoan population of the soil; *Phil Trans. Roy. Soc. Lond* 211B 317-350
  16. Dangeard P 1900 Etude de la karyokinise chez l' *Amoeba hyalina* sp. nov., *Botaniste* 7 Ser., pp. 49-83
  17. Das S R, Willaert E and Jadin J B 1974 Studies on mitotic division in *Naegleria jadini*; *Ann. Soc. belge Med. trop.* 54 141-46
  18. Dobell C 1914 Cytological studies on three species of *Amoeba*—*A. iacertae* Hartmann, *A. glebae* n. sp. *A. fluvialis* n. sp. *Arch. Protistenk.* 34 139-190
  19. Douglas M 1930 Notes on the classification of the amoeba found by Castellani in cultures of yeastlike fungus; *J. Trop. Med. London* 33 258-259

20. Duma R J 1972 primary amoebic meningoencephalitis. *Critt. Rev. Clin. Lab. Sci.* **3** 163-192  
—, Helwig W B and Martinez A J 1978 Meningoencephalitis and brain abscess due to a free-living amoeba; *Ann intern. Med.* **88** 468-73
21. Ferrante A and Rowan-Kelly B 1983 Activation of the alternative pathway of complement by *Acanthamoeba culbertsoni*; *Clin exp Immunol.* **54** 477-485
22. Fowler M and Carter R F 1965 Acute pyrogenic meningitis probably due to *Acanthamoeba* sp., a preliminary report; *Brit. Med. J.* **2** 740-742
23. Fulton C and Guerrini A M 1969 Mitotic synchrony in *Naegleria* amoebae; *Exptl Cell Res* **56** 194-200
24. Gogate Alka, Singh B N, Deodhar Leena P and Jhala H I 1984 Primary amoebic meningo-encephalitis caused by *Acanthamoeba*, *J. Postgraduate Medicine.* **30** 125-28
25. Gordeeva L M 1973 Isolation and cultivation of *limax* amoebae capable to grow at 37°C, *Prog Protozool* (Clermont-Ferrand) p. 159
26. Griffin J L 1978 pathogenic free-living amoebae; in *Parasitic Protozoa* Vol. 2, pp 507-549 ed. J P Kreier (New York, San Francisco & London: Academic Press)  
— 1983 The Pathogenic amoeboflagellate *Naegleria fowleri* environmental isolates, competitors, ecologic interactions and the flagellate-empty habitat hypothesis. *J. Protozool* **30** 403-9
27. Holbrook T W, Boackle R J, Parker P W and Vesely, J 1980 Activation of the alternative complement pathway by *Naegleria fowleri*; *Infect. Immun* **30** 58-61
28. Ivanic M 1926 Zur Kenntnis der Entwicklungsgeschichte einer neuen *Hartmannella*-Art (*H. testudinis* spec. nov.); *Zool Anz.* **68** 87-95
29. Jadin J B 1973 De la méningo-éncéphalite amibienne et du pouvoir pathogène des amibes (*limax*); *Ann. Biol* **12** 305-42
30. John D T 1982 Primary amoebic meningo-encephalitis and the biology of *Naegleria fowleri*; *Ann. Rev. Microbiol.* **36** 101-23
31. Kudo R R 1954 *Protozoology* 4th ed. (Springfield, Illinois, USA: Charles C. Thomas)
32. Ma P, Wilaert E, Juechter K B and Stevens A R 1981 A case of keratitis due to *Acanthamoeba* in New York, and features of 10 cases; *J. Infect. Dis.* **143** 662-67
33. Martinez A J 1976 Amoebic meningo-encephalitis due to *Naegleria* and *Acanthamoeba*, Clinico-neuropathological correlation; *Proc. Int. Confr. on Amoebiasis. Mexico city, Oct. 27-29, 1975* (Mexico Instituto del seguro Social) pp. 64-81  
— 1980/Is *Acanthamoeba* encephalitis an opportunistic infection? *Neurology* **30** 567-574

- 1982 Acanthamoebiasis and immunosuppression, case report; *J Neuropath. Exp Neurol.* 41 548-557
- 1983 Free-living amoebae: pathogenic aspects (A review) *Protozool. Abstracts* 7 293-306
- and De Jonckheere J F 1981 Les infections par les amibes libres; *Bull. l'Institut Pasteur* 79 171-205
- , Stelo-Avila C, Alcala H and Willaert E 1980 Granulomatous encephalomyelitis, Intracranial arteritis and mycotic aneurysm due to a free-living amoeba; *Acta neuropath (Berlin)* 49 7-12
- 34. Misra R and Sharma A K 1980 Pathogenicity in mice of strains of *Acanthamoeba culbertsoni* (Singh & Das, 1970) and *A. rhysodes* (Singh, 1952) from Indian soils and the present status of the genera *Hartmannella* and *Acanthamoeba*, *Indian J Parasitol* 4 135-144
- 35. Neff R J 1957 Purification, axenic cultivation and description of a soil amoeba, *Acanthamoeba* sp., *J Protozool* 4 176-182
- 36. Page F C 1967a Taxonomic criteria for *limax* amoebae, with description of three new species of *Hartmannella* and three of *Vahlkampfia*; *J Protozool* 14 499-520
- 1967b Redefinition of the genus *Acanthamoeba* with description of three species; *J Protozool* 14 709-724
- 1968 Generic criteria for *Flabellula*, *Rugipes* and *Haylodiscus* with description of species, *J. Protozool.* 15 9-26
- 1974 A further study of taxonomic criteria for *limax* amoebae, with description of new species and a key to genera; *Arch Protistenk* 116 148-188
- 1976a A revised classification of the *Gymnamoeba* (Protozoa: Sarcodina); *Zool J Linn Soc.* 58 61-77
- 1976b *An Illustrated Key to Freshwater and Soil amoebae* (UK: Freshwater Biological Association, Scientific Publication No. 34)
- 1981 A light-and electron microscopical study of *Protacanthamoeba caledonica* n. sp., type species of *Protacanthamoeba* n.g. (Amoebida, Acanthamoebidae); *J Protozool.* 28 70-78
- 37. Pan N R and Ghosh T N 1971 primary amoebic meningo-encephalitis in two children; *J. Indian Med. Ass.* 56 134-137
- 38. Patras D and Andujar J J 1966 Meningo-encephalitis due to *Hartmannella* (*Acanthamoeba*); *Amer. J. Clin. Pathol.* 46 226-233
- 39. Pussard M 1966 Le genre *Acanthamoeba* Volonsky, 1931 (*Hartmannellidae* Amoebida); *Protistologica* 2 71-93
- 40. Rafilko J S 1947 Cytological observations on the amoeboflagellate, *Naegleria gruberi* (Protozoa); *J. Morph.* 81 1-44

41. Ray D L and Hayes R E 1954 *Hartmannella astronyxis*: A new species of free-living amoeba. Cytology and life cycle; *J. Morph.* **95** 159-188
42. Rowan-Kelly B, Ferrante A and Thong Y H 1980 Activation of complement by *Naegleria*. *Trans. Roy. Soc. Trop. Med. Hyg.* **74** 333-335
43. Russel E J and Hutchinson H B 1909 The effect of partial sterilization of soil on the production of plant food; *J. Agric. Sci.* **3** 111-144.
44. Sandon H 1927 *The Composition and Distribution of the Protozoan Fauna of the Soil* (Edinburgh & London : Oliver and Boyd)
45. Sawyer T K and Griffin J L 1975 A proposed new family, Acanthamoebidae n. fam. (order Amoebida), for certain cystforming filosea amoebae; *Trans. Amer. Micros. Soc.* **94** 93-98
46. Schardinger F 1899 Entwicklungskreis einer *Amoeba lobosa* (Gymnamoeba) : *Amoeba gruberi*; *S. K. Acad. Wiss. Wien* **108** 713-734
47. Schuster F L 1979 Small amebas and amoeboflagellates; in *Biochemistry and Physiology of Protozoa* 2nd ed. Vol. 1 (New York, London, Toronto, Sidney and San Francisco: Academic Press) pp. 215-285
48. Singh B N 1941 Selectivity in bacterial food by soil amoebae in pure mixed culture and in sterilised soil; *Ann. Appl. Biol.* **28** 52-64
  - 1942 Selection of bacterial food by soil flagellates and amoebae; *Ann. Appl. Biol.* **29** 18-22
  - 1945 The selection of bacterial food by soil amoebae and the toxic effects of bacterial pigments and other products on soil protozoa; *Brit. J. Exp. Path.* **26** 316-325
  - 1946 A method for estimating the numbers of soil protozoa, especially amoebae, based on their differential feeding on bacteria; *Ann. Appl. Biol.* **33** 112-119
  - 1952 Nuclear division in nine species of small freeliving amoebae and its bearing on the classification of the order Amoebida; *Phil. Trans. Roy. Soc. Lond.* **B236** 405-461
  - 1955 Culturing soil protozoa and estimating their numbers in soil; in *Soil Zoology* (London : Butterworth's Scientific Publications) pp 403-11
  - 1960 Interrelationship between micro-predators and bacteria in soil; *Proc. 47th Indian Sci. Congr. Asso.*, Part 11 145-158
  - 1963 Recent advances in soil protozoology; in *Proc. 1st Summer School Zoology* (Delhi : Manager of Publications, Government of India) pp 79-103
  - 1972 Classification of amoebae belonging to the order Amoebida with special reference to the pathogenic free-living forms; *Current Sci.* **41** 395-403
  - 1973 Current status of the problem of exogenous and endogenous amoebiasis; *J. Sci. Industr. Res.* **32** 399-432

- 1975 *Pathogenic and Non-pathogenic Amobae*. (London & Basingstoke: The Macmillan Press Ltd.)
- 1981 Nuclear division as the basis for possible phylogenic classification of the order Amoebida Kent, 1880; *Indian J. Parasitol.* **5** 133-153
- and Crump L M 1953 The effect of partial sterilization by steam and formalin on the numbers of amoebae in field soil; *J. gen. Microbiol.* **8** 421-426
- and Das S R 1970 Studies on pathogenic and non-pathogenic small free-living amoebae and the bearing on nuclear division on the classification of the order Amoebida, *Phil. Trans Roy. Soc. Lond.* **B259** 435-476
- and — 1972a Occurrence of pathogenic *Naegleria aerobia*, *Hartmannella culbertsoni* and *H. rhyodes* in sewage samples of Lucknow; *Current Sci.* **41** 277-281
- and — 1972b Intranasal infection of mice with flagellate stage of *Naegleria aerobia* and its bearing on the epidemiology of human meningo encephalitis; *Current Sci.* **41** 625-628
- and Dutta G D P 1984 Small free-living amoebae : Soil as a suitable habitat, isolation, culture, classification, pathogenicity, epidemiology and chemotherapy; *Indian J Parasitol* **8** 1-23
- and Hanumaiah V 1979 Studies on pathogenic and non-pathogenic amoebae and the bearing of nuclear division and locomotive form and behaviour on the classification of the order Amoebida. Monograph No. 1 of the Association of Microbiologists of India, *Indian J Microbiol* pp 1-80
- and Sharma A K 1983 Antigenic relationship among seven species of *Acanthamoeba* using immobilization and gel diffusion precipitin tests; *Indian J. Parasitol.* **7** 149-163
- 49. Singh L M and Singh B N 1966 Antigenic difference in free living and parasitic amoebae as revealed by immobilization reaction; *Indian J. Exp Biol.* **4** 49-50
- 50. Vahlkampf E 1905 Beitrage zur Biologie und Entwicklungsgeschichte von *Amoeba limax*, U.S.W.; *Arch. Protistenk* **5** 167-220
- 51. Visvesvara G S and Balamuth W 1975 Comparative studies on related free-living and pathogenic amoebae with special reference to *Acanthamoeba*; *J. Protozool.* **22** 245-256
- 52. Visvesvara G S and Healy G R 1975 Comparative antigenic analysis of pathogenic and free-living *Naegleria* species by gel diffusion and immuno-electrophoresis techniques; *Infect. Immun.* **11** 95-108
- 53. Volkonsky M 1931 *Hartmannella castellanu* Douglas et classification des *Hartmannelles*, *Arch. Zool Exp. Gen* **72** 317-339
- 54. Willaert E, Stevens A R and Healy G R 1978 Retrospective identification of *Acanthamoeba culbertsoni* in a case of amoebic meningoencephalitis; *J. Clin. Pathol.* **31** 717-720





**Triloki Nath Khoshoo** (b. 7 April 1927) did Ph.D. (1958) from Panjab University, Chandigarh. He was CSIR Distinguished Scientist (1986-90); Secretary, Government of India (1982- 85); Director, National Botanical Research Institute, Lucknow (1976-82).

Khoshoo has made significant contributions to the genetic-evolutionary understanding of several non-agricultural economic plants (including ornamentals). He applied this knowledge for their improvement and has evolved over 30 cultivars of ornamentals, many of which are in nursery trade.

He has elucidated the genetic system of gymnosperms (softwoods in particular). Many of the basic concepts he has developed are known for their originality and have stood the test of time. He is author of 3 books and editor of 9 books.

Khoshoo is Fellow of Indian Academy of Sciences, Indian Society of Genetics and Plant Breeding, Indian Botanical Society, National Academy of Agricultural Sciences (India), and Institution of Engineers, President, Bioenergy Society of India (1985-86), National Academy of Sciences, India (1985-86, also Honorary Fellow), Indian Society of Genetics and Plant Breeding (1986), and Indian Society of Tree Scientists (1988-1991); Member Council (1983-84), Foreign Secretary (1985-88), INSA. He is the recipient of Rafi Ahmed Kidwai Medal and Prize (1977); UGC National Lecturer (1981), G. P. Chatterjee Lecture Award (INSA) (1985); Birbal Sahni Gold Medal (1982), Seth Memorial Medal (1983), Ramdeo Misra Medal (1984); Dayawati Vira Medal (1985); Sanjay Gandhi Award (1986); President, Indian Science Congress (1986), Distinguished Service Award (Indian Science Congress) (1988); Om Prakash Bhasin Foundation Award (1989); Distinguished Service Award (Indian Society of Genetics and Plant Breeding) (1991); Jawaharlal Nehru Fellow (1991-93); Felicitation by Bochasanwasi Shri Aksharapurushotam Sanstha (Ahmedabad) (1992), Padma Bhushan (1992).

*Triloki Nath Khoshoo was elected to the fellowship of the Academy in 1977*

**Address:** C/o TERI House, 9 Jor Bagh, New Delhi 110 003; Tel: (011) 617025, 4622246, 4623935; Telex: 31-61593; Fax: 91- 11-4621770; Res: 103, Pocket H, Sarita Vihar New Delhi 110 044; Tel: (011) 6849625.

## BIOMASS-ENERGY: RETROSPECT AND PROSPECT (EXTENDED SUMMARY)

T N KHOSHOO

*The goal and role of biomass energy in the Indian context, particularly for rural poor, where firewood is still the principal form of energy, is discussed. It brings out the immediate research needs in firewood forestry which can be taken up on marginal land to bridge the widening gap between supply and demand of firewood. There is also need to undertake R and D on agricultural alcohol, hydrocarbon plants, algae, biological hydrogen, etc. Most of these are in urgent need of domestication and genetic upgradation. Biomass cannot supply all our energy needs, but will supplement the same. There is a nexus between food, firewood and fodder as reflected by the three land-based vocations, agriculture, forestry and animal husbandry. There are many social, economic and environmental advantages of such an agri-silvi-pastord system.*

Ever since the human being discovered fire some 150,000 years ago, firewood provided him with warmth when it was cold, protection from predators, more digestible food, higher survival rate and colonisation of habitats which were inhospitable earlier. Firewood soon became the principal source of energy and even today over 50% of wood is put to its original use, i.e. as fuel. Firewood and charcoal gave way to coal, coal to electricity, and since 1859 oil has been the major source of energy.

Two important points emerged from the First Energy Crisis in 1973. Firstly, all nations suddenly became aware of their total dependence on only one form of energy and realized that they were, in fact, living in a Petroleum Society. Secondly, it became clear that such a crisis would not have affected any nation if a broad energy policy, involving many sources, had been adopted. There then followed the Second Energy Crisis due to scarcity of firewood for cooking and heating. This crisis has been more serious, as it affects a very large rural population in the entire developing world. In India, 68.5% of the energy used in households is in the form of firewood and 64.2% of it is collected from natural sources. For a variety of reasons, wood is still the principal source of energy in rural areas and nearly half of the world's wood production goes into its original use, i.e. cooking and heating. However, on account of population pressure, the demand for firewood has out-stripped natural regeneration and planting, so

much so that in some areas, though there is food to eat, not enough wood is available to cook it. At present, there is a big gap between demand and supply of firewood resulting in depletion of the forest cover, which, in turn, has proved to be ecologically disastrous, as it leads to floods, soil erosion and, ultimately, shortage of water.

One of the important methods of insulating the country against impending disaster is to take to the photosynthetic model of development through large-scale use of biomass, particularly firewood. Such an option, apart from meeting energy needs, would help restore the relationship between humankind and its environment. Therefore, a very massive tree plantation programme is necessary both for energy needs and overall eco-development.

In India, the shortfall in fire wood production by 2000 AD has been estimated to be of the order of 137 million tonnes, which would require 34 million hectares of land, at about four oven-dry tonnes of wood per ha/year, and a minimum annual outlay of Rs. 500 crore for 17 years, to be made good. Once such a programme is successful, dung and plant-based residues would become available as organic fertiliser and industrial feedstock.

Among the renewable alternatives, solar energy, captured by the plants through the process of photosynthesis, is the most important, especially because photosynthesis is the key process in the life-support system of this planet. Furthermore, the plant-based energy systems are not only renewable but they also remove carbon dioxide from the atmosphere before turning it back, with no overall quantitative increase: thus; they help to contain environmental pollution. Photosynthesis is at the base of all biomass production and food chains. It converts physical energy into chemical energy and generates oxygen, the life sustaining gas. Whereas net photosynthesis uses only 0.1% of sunlight, it produces organic matter only 10% of which accounts for the total energy used by humankind and only 0.5% accounts for the entire food requirement of the human race. Increase in photosynthetic efficiency automatically increases production of organic matter. Various types of biomass and/or biofuels available are: firewood, agricultural alcohols, vegetable oils, hydro-carbon plants, particularly those yielding rubber and petroleum-like materials; fresh weeds, sewage-grown algae, algal hydro-carbons, and biologically produced hydrogen using halobacteria, algae, azolla and even higher plants. Every feedstock or bioconversion process has its own merits and

demerits, and the different routes available to generate solid, liquid and gaseous fuels are: anaerobic and enzymatic digestion and thermochemical conversion.

The basic premise for raising firewood plantations is that firewood needs cannot and should not be met from the natural forest stands, as it would reduce that forest area, which is now only 12% in India, although 33% is envisaged under the National Forest Policy Resolution (1952) of the Government of India. The destruction of our forests will not stop as long as fuel for cooking is either not provided free to the rural poor, or is not made very cheap. In fact, the time has come when we must plan for meeting timber and wood needs for industrial purposes from plantations captive to industry and not from natural forests. In order to obtain the maximum productivity of firewood in the minimum time, high-density and short-rotation biomass of fast and hardy species of deciduous trees and shrubs should be grown exclusively on non-agricultural or marginal lands, so as to avoid competition with food crops. The yields of firewood reported so far are indeed over-optimistic and there is a need to conduct meticulous scientific experiments, because estimated yields are a very crucial factor in planning firewood plantations for domestic use and for generation of electricity. The numerous advantages of large-scale firewood plantations range from the strictly local to environmental and social imperatives.

Another dimension to the problem is that the degraded lands proposed to be used for biomass production are at present used, free of cost, for grazing. If used for firewood biomass, such lands will have to be closed to grazing immediately. To meet the situation thus created, either tree species with a fodder value will have to be grown or fodder legumens and/or grasses will have to be produced in conjunction with firewood species.

The immediate research needs are:

To select tree/shrub species for wide genetic variation and adaptability to marginal or low nutrient soils. Thus, the selected species should be hardy and require low inputs of water, fertilizer and plant protective measures.

To choose species with multiple uses, like fuel, fodder, fertilizer and fibre and high regenerative potential and/or coppicing ability, without loss of vigour under conditions of competition; with

minimum amount of bark wood with high calorific value and capability to burn without sparks and toxic smoke, etc.

- Nitrogen-fixing capacity of the species would be an additional advantage, and other methods of biological soil fertilization for deficient soils need to be worked out in view of the fact that high density and short rotation will cause a very high drain of nutrients from soil with hardly any litter fall available for recycling.
- Agro-technological package of cultural practices for individual species and specific habitats needs to be looked into in combination with appropriate fodder legumes and/or grasses.
- Standardization of tissue culture techniques for production, on a mass scale in a very short time, of elite clones will be required to meet a very large demand for planting material (10,000 and above per hectare).
- Germ plasm collection of all the relevant species and their variants will have to be made for purposes of location-specific adaptability trials as also for incorporation in breeding programmes.
- Breeding procedure for genetic upgrading for the parameters listed above will need standardisation.

Agricultural alcohol, though a successful transport fuel in Brazil, is not relevant to India, as it may be more economical to use it as an industrial feedstock. Furthermore, it may not be wise to displace sugarcane and/or tapioca as food crops. Sugarcane, besides, is now a 'political' crop. However, the future feedstock for alcohol is lignocellulose and research and development (R&D) work has to be intensified to make it a commercial raw material. Similarly, although vegetable oils are a promising fuel for diesel engines, due to acute shortage of edible oils, as also their higher price as compared to petroleum in India, these oils do not have an immediate future as fuel.

The new hydrocarbon plants, such as guayule and euphorbias, are hardy and most relevant to India and can be grown in arid regions. The demand for natural rubber is unabated, because, it is preferred to synthetic elastomers; besides, the latter are based on petroleum feedstock. To meet a shortfall of 65,000 tonnes of natural rubber, guayule needs to be raised in over 26,000 ha in arid zones by 1985.

Among the aquatic biomass, water weeds, such as water hyacinth, common in water bodies, are already being used. Sewage, though an environmental hazard, can be a source of biogas as also a culture medium for single cell protein for poultry. During the process, sewage itself is rendered benign.

Algae are a very promising source of hydrocarbons and deserve attention for production and processing by catalytic hydrogenation. The whole area of marine biomass has been badly neglected and requires examination in view of the fact that the sea constitutes about 71% of earth's surface.

Production of biological hydrogen through bio-photolysis, as also bio-electricity, involves the use of hydrogenase and halobacteria with purple membranes, respectively. Both of these merit serious consideration and have a long-term potential.

There is an urgent need to domesticate as also genetically upgrade all the new and upcoming energy crops, microbes and algae, which are at present wild and have low productivity per unit area, per unit time. This can be possible if there is a heavy input of high science. The area of biomass energy is indeed multi-disciplinary and requires R & D in subjects such as botany, phytochemistry, microbiology, agriculture, fuel research, heat physics, different branches of engineering, etc., on plants as alternative sources of diverse types of energy.

In general, energy requirements of the rural poor, who often have food but not enough fuel to cook it, have been neglected so far. In view of the acute shortage of firewood, they perforce poach on forests, plantations, etc. It is, therefore, essential that, like other basic needs, such as food, shelter, clothing and medicare, energy for cooking and heating should also be considered a basic need and guaranteed by the state. This has become all the more necessary on account of the very wide regional disparity in energy use. In fact, affluent nations are using energy in such a reckless manner as to lead to a highly disproportionate use of this finite resource and increased pollution. One way to meet this situation is to follow a broad energy policy in which many forms of energy are utilised, each supplying not more than 20-25% of the total energy requirement. In a country like India, which has achieved food self-sufficiency to an extent, but is deficient in energy, the non-agricultural land will tend to be used for energy crops, in contrast with those countries which are deficient in both energy and food, and have little or no land to spare.

Obviously, biomass, agriculture and animal husbandry form an integrated and symbiotic system. To meet the energy crisis, and the increased demands for food, fibre and feed, the only option for India is to widen her agro-forestry base. Since agriculture is indeed a dynamic, living and continuing system, the role of agro-forestry is to maintain land in a living and productive form, so that human life is sustained for a very long time to come. While agricultural waste residues are a valuable feedstock for bio-fuels which can be utilised in a number of ways for productive purposes, it is necessary to collect all the available waste residues to enable natural soil enrichment to continue through degradation by soil micro-organisms.

Although even enlightened people think that wood is a fuel of the past, it has been gaining in importance steadily as a source of energy in developed countries particularly, important as it is for the developing ones. Since there are several economic, social and environmental advantages accruing from biomass, it must receive greater attention of policy-makers, planners and funding agencies in the developing countries. Although biomass is a dire necessity for the developing world, there is greater awareness about its importance and R & D in the developed countries. Obviously, in the times to come, the latter would make far greater use of biomass than the developing world.

While biomass goes well with the social, cultural and economic milieu of the country, there is a wide difference between the perception of a villager and a national planner about the crisis: the former sees it as a local problem and feels that wood is a gift of God and he has a right to it even by poaching; the latter sees it as a problem of denudation of forests, leading to ecological disaster. However, for the success of the programmes of firewood plantations, the most critical factor is the villager who is both their guardian and end-user. Unless he is made aware of the dire implications of destruction of forests and plant cover in general and the fact that it is in his own self-interest to preserve and make rational use of the plantations, poaching will never stop. Associated with this is the need to evolve and popularise cooking stoves which are energy efficient and smokeless. The smoke itself sometimes contains materials which may be toxic. There has to be a mass awareness programme encompassing all aspects, with people's participation. For such participatory forestry programmes, apart from the forest departments there is need for a cadre of trained 'barefoot foresters' who can work with the people.

Biomass may not be the panacea for all our energy problems, but it will, no doubt, help to reduce substantially our dependence on fossil fuels. Being socially and environmentally relevant, biomass enables us to keep our air, water and land clean and to manage our life support system in a sustained manner. Two things are needed: (i) India being a predominantly agricultural country a perceptible tilt in favour of plants and plant sciences in our planning process by adoption of the Photosynthetic Model of Development, and (ii) India has to be made increasingly greener. Among other things, such a model envisages revegetating the uncultivated half of India to make the country fresh and verdant. This would have distinct environmental, social and economic benefits and will help in the following ways:

- Conservation and improvement of soil and water by reduction of surface run-off, nutrient leaching and soil erosion, and increasing soil nutrients by addition and decomposition of litter fall; abatement of dust pollution;
- Stabilisation of catchments and watersheds;
- Control of floods;
- Better microclimate by decrease in soil surface temperature and decline in evaporation of soil moisture on account of mulching and shading;
- Conservation of biological diversity;
- Resolving energy crisis in a decentralized manner;
- Reduction of pressure on forests;
- Employment generation;
- Creation of aesthetic and pleasing landscapes;
- Better health;
- Better quality of life;
- Halting influx of rural population into urban areas; and
- Decentralised economy.

The only way to restore the forest cover is to take tree and fodder planting programmes on a priority basis under the National Rural Employment Programme (NREP) and/or the recently announced Employment Guarantee Scheme (EGS) of the Government of India. In



fact, the people involved in such programmes could be organised in a Conservation Corps. Meaningful results can be obtained only if tree and fodder planting is taken up on a war footing and work started as expeditiously as possible. Even if this programme is started today, the first results would not be discernible before five years. Besides, this would instil in the young people a work ethic and a sense of pride.

The advantages of producing and utilising biomass in the social, cultural and economic milieu of India may now be enumerated.

- It is essential a decentralised energy system, being basically local and utilising locally-produced materials. Thus, it is immune to international political pressures;
- Biomass offers clean fuel/energy and keeps the environment more or less clean, there being no significant environmental pollution as compared to large-scale use of oil, coal and nuclear power. Biomass also helps to correct eco-imbalances;
- Due to fossil fuel energy-based systems, CO<sub>2</sub> content of the atmosphere has increased nearly 10 times during the past 100 years. Burning biomass, or fuels derived from it, does not alter CO<sub>2</sub> concentration in the atmosphere; in fact, it helps to recycle CO<sub>2</sub>. The importance of photosynthesis, which is the basic process involved in the production of biomass, can be gauged from the fact that every year 300 billion tonnes of carbon are fixed by this process in the form of terrestrial biomass and it stores 10 times more energy than the world consumes in a year;
- Biomass provides besides fuel, a wide range of feedstocks for the organic chemicals industry, which, at present, is mostly petroleum-based;
- Bioconversion of biomass is less energy intensive and involves high sciences, but not necessarily high and capital-intensive technology;
- Biomass will utilise marginal, arid/semiarid and uncultivated land at present lying fallow;
- Production of biomass on such land will improve soil and its water retention capacity and there is likely to be reduction of erosion by wind and water; and
- Biomass is socially relevant, being labour-intensive, with opportunities for generation of employment/income for the poorer

sections. Furthermore, the quality of life in rural areas would improve and women and children, who spend most of their time in gathering firewood, would use their time more profitably, such as for their education. It may also help in halting or reducing the rate of migration from rural to urban areas. The overriding consideration is the provision of energy to a large number of poor people very quickly and inexpensively.

Before considering one of the major disadvantages of utilising agricultural and forest residues for fuel purposes it may be pointed out that, while a good portion of crop plants is harvested, the roots and some of the above ground portions as also the leaves are left in the field. Similarly, in the forests, a lot of plant residues, roots, bark, branches and leaves, flowers and fruits that fall from the trees, remain on the ground. All crop and forest residues contain considerable proportion of nutrient elements absorbed from the soil during the growth of the plants. All nutrients in the residues are recycled to enrich the soil. In this way, a sizeable part of organic matter fixed each year is returned to the soil; besides, humus is also formed. There is, thus, a biological turnover in the soil and the cycle operates continuously, depending upon the availability of the substrates (residues). Thus, the nutrients are released in soil through a highly efficient decomposer chain, depending upon biochemical composition of litter, microbial flora and fauna inhabiting the soil and hydro-thermal conditions of the soil. However, with the removal of the residues from the field and forest, the soil would become depauperate, as there would be depletion of nutrients and lack of humus formation as also the conditioning process of the soil. Because of chronic shortage of fertilizer and other agricultural chemicals, it will not be advisable to remove the entire crop and forest residues for utilisation as feedstocks for energy. In addition to the foregoing constraint, the collection and transport of residues, which are dispersed over a large area, to the utility sites for processing, would reduce energy gains.

Although biomass may not always be profitable on techno-economic considerations, this energy is most relevant, both socially and environmentally, due to the many intangible benefits enumerated above. The techno-economics will change materially as soon as productivity per unit area, per unit time is increased by improving plants and microbes involved in biomass production and conversion through induction of genetical and agronomical research and developmental components. One way to reduce the cost of biomass-fuels is to aim at total utilisation of

biomass, e.g. rice husk not only to be used as fuel, but also as a source of a number of chemicals for a variety of uses.

### RESEARCH NEEDS

Oil was discovered in 1558 and was initially valued only for kerosene, for lighting purposes. Later were developed from it natural gas, liquified petroleum gas, gasoline, fuel oils, lube oils, asphalt, paraffin, etc. By 1975, a stage of complete utilisation of a barrel of oil was reached. Now that oil is fast running out, biomass is regarded important among the promising alternatives, but complete utilisation of a tree to meet the requirements of sugar, natural and synthetic fibres, lumber, wood chemicals (like rubber, turpentine, lignin, phenols, etc.), synthetic chemicals and fuels (methanol, methane, ethanol, ethane, gasoline), etc., has not been possible so far. Technologies to accomplish the same are expected to be developed by AD 2025. Here then lies a challenge for scientists, and a radical research and development approach is needed not only for conversion and utilisation of biomass, but also for its production. The latter, among other things, requires improvement in photosynthetic efficiency. Plants with a high energy potential can be cloned and grown in high plant densities as feedstocks for cleaner fuels and energy.

As stated earlier, the R and D on energy from biomass is at a fairly advanced stage in the developed countries, because they have the requisite wide scientific base, capital and manufacturing capability for its complete utilisation. Obviously, they would reap the benefits of such researches much earlier than the developing nations for whom it is indeed a dire necessity.

Keeping in view some of the foregoing points, the immediate research needs may be enumerated as under:

- While data on commercial energy are reasonably reliable, those on biomass are rather speculative. There is, thus, a need for organisation of a database regarding location-wise pattern of biomass availability and consumption. This also includes availability and use of firewood and agricultural residues/wastes, animal dung, etc., together with information on specific energy plants;
- Before embarking on the "fuel from crop and forest residues programme", it is necessary that critical investigations are carried out on the possible consequences of biomass removal, from

agricultural fields and forests in respect of the possible depauperization of soil and degradation of environment. The amount of biomass to be left in the fields and in forests for enrichment of soil by natural processes needs to be estimated, location-and species-wise;

- A R and D programme on identification and development of silvicultural, agronomic and management practices for maximising biomass production under cultivated conditions is urgently called for;
- First-tier performance trails in specific locations with specific plant species, having multiple uses, need to be undertaken as a prelude to massive afforestation/plantations for fire/fuelwood on non-agricultural land;
- Methods need to be evolved to compensate for the faster depletion of nutrients from soil due to short rotation and high density plantations;
- Research on the utilisation of agricultural residues as also aquatic biomass and municipal wastes for biogas generation requires to be undertaken;
- Research on bioconversion of ligno-cellulose needs to be taken up;
- The micro-organisms involved in bioconversion as also energy plants needs to be genetically upgraded for high productivity and photosynthetic efficiency;
- Thermochemical processes/technologies need improvement for getting higher yield of products;
- The energy and non-energy uses of biomass need integration and byproduct utilisation requires improvement so as to enable utilisation of complete biomass;
- Wood burning stoves need improvement for attaining higher heat efficiency;
- A trained cadre of scientific and technical personnel in the area of production and utilisation of biomass needs to be created in order to enhance capabilities in this area; and
- Demonstration and awareness programmes need to be organised:

Production of energy from biomass being a multidisciplinary area, there is an urgent need for cooperation among concerned R and D agencies and departments of environment, science and technology, energy, forestry, agriculture and community development. There is also need for circumscription of the role of biomass energy for continued growth so as to achieve better balance among rural, urban agricultural, forestry and industrial needs. It may, however, be emphasized that biomass can provide most of the products now obtained from oil and, in time to come, cellulosic materials will be increasingly utilised as feedstock for this purpose.

In the end, let us hope that R and D on biomass ushers in a new society in our country, which will rely more on renewable energy, sources, such as biomass and solar technology, with smaller and more widespread industrial establishments that utilise biomass-based fuel/energy systems. This will help us to manage our life support system better, so that existence of humankind as also associated ventures such as agriculture, forestry and the relevant industry are not threatened. To achieve this, India will have to be made greener in years to come.

The domestic energy problems, particularly of the rural community and the urban poor, in a developing country are, indeed, very complex. They are closely linked with poverty and inequality. There are no magic solutions, but given the sustained commitment of the people, they are not insurmountable. If we do not give the highest priority to these questions now, the price we will have to pay in future will be colossal—ecological disaster and an economic crisis. Let us remember that good ecology is good economics.



**Prem Narain** (b. 3 January 1934) did D.Sc. (1984) from University of Edinburgh, UK. He is Principal Scientist, Indian Agricultural Research Institute, New Delhi (1992- ); and was formerly Director (1981-92), Joint Director (1978-81), Senior Professor (1970-78), Indian Agricultural Statistics Research Institute, New Delhi.

Narain's researches have opened up new vistas in the theory of genetics and their application to the breeding of domesticated plants and animals. He has also contributed to the theory and application of studies to crops, livestock and fisheries. He was

instrumental in providing efficient statistical and computer techniques for agricultural research and development. Also developed a blueprint of KRISHINET — a nationwide agricultural research information system. He has authored a book 'Statistical Genetics' (Wiley Eastern, 1990)

Narain is Fellow of Indian Academy of Sciences, National Academy of Sciences (India) and National Academy of Agricultural Sciences (India) (currently Treasurer), Royal Statistical Society (UK); Council Member, International Biometric Society, USA (1988-91); President, Indian Society of Agricultural Science; Secretary-General, Federation of Indian Societies of Agricultural Sciences and Technology; Secretary, Indian Society of Agricultural Statistics. He is the recipient of Rafi Ahmed Kidwai Memorial Prize (1977); G.P. Chatterji Memorial Lecture Award (INSA) (1987); Sankhyiki Bhushan Award (Indian Society of Agricultural Statistics) (1991); Om Prakash Bhasin Foundation Award (1992).

*Prem Narain was elected to the fellowship of the Academy in 1978.*

**Address:** Principal Scientist, Indian Agricultural Research Institute, Pusa, New Delhi 110 012; Tel: (011) 5782817; Telex: 031-77161 IARI-IN; Fax: 91-11-5740722; Res: B-5 IARI Campus, New Delhi 110 012; Tel: (011) 5788770.

# THE EVOLUTIONARY DYNAMICS OF QUANTITATIVE CHARACTERS

P NARAIN

## INTRODUCTION

The fact that evolution occurs is well recognised but what causes it to occur has been a matter of debate since the time Darwin<sup>6</sup> ascribed it to 'Natural Selection'. With continued natural selection for several generations, individuals possessing favourable characteristics have an advantage over those lacking them and as such they pass on such characteristics to their offspring for further modification and adaptation. Over the evolutionary time scale, this results in modified descendants of different ancestors living in geological times. However, for such a process to occur, it is necessary to have heritable variation and to understand factors governing it. Darwin unfortunately failed to discover the precise nature of such variation. It was only when Mendel's laws were re-discovered in 1900, that the precise nature of such variation, in the form of gene mutation as raw material for evolution to act upon, was properly understood. However, since the characteristics favoured during adaptive evolution are primarily quantitative in nature whereas mutation produces discontinuous variation, the compatibility between Darwinism and Mendelism was hotly debated for quite some time. Fisher<sup>7</sup> paved the way for a reconciliation between the two by showing how the observed correlation between relatives for such characteristics can be explained entirely in terms of the effects produced by Mendelian genes. Not only that, Fisher<sup>9</sup> further showed that the rate and direction of evolution are primarily determined by natural selection. According to his Fundamental Theorem of Natural Selection, the rate of increase in the average fitness of a population is proportional to additive genetic variance in fitness. Fisher<sup>8,9</sup> also investigated, for the first time, the problem of the maintenance of genetic variability in natural populations for quantitative characters and gave a model in which fitness was assumed to decrease in proportion to the squared deviation from the optimum. In this he initiated investigation on a problem which was not only followed by several workers subsequently but is still being vigorously pursued since the mechanism for the maintenance of such variability in natural populations is far from being well understood.

Subsequent to Fisher's work, Wright<sup>30,31</sup>, Haldane<sup>13</sup>, Robertson<sup>26</sup>, Latter<sup>22</sup> and Bulmer<sup>2,3</sup> dealt with models based on multiple loci but each with only two segregating alleles. Modern understanding of molecular genetics, however, indicates that a gene is subdivisible into a very large number of variable nucleotide sites so that a model involving segregation of infinitely many alleles at a given locus is more realistic than the early attempts. Crow and Kimura<sup>5</sup> introduced, for the first time, an infinite allele model which was later studied by Kimura<sup>14</sup> in detail. In this study, mutation was considered as producing multiple alleles with varying phenotypic effects in a continuous manner at each of the several loci involved in the inheritance of quantitative characters. Later on, Latter<sup>23</sup> adopted such a model in a discrete-time fashion. By extending this model to multiple loci, Lande<sup>17-21</sup> considered the effects of linkage and linkage disequilibrium. Fleming<sup>10</sup> studied extensively the entire equilibrium structure of the multi-allele multi-locus case, both for discrete as well as continuous time situations.

The investigations of Latter<sup>22</sup> and Bulmer<sup>2,3</sup> based on diallelic loci, led to the conclusion that the equilibrium genetic variance was independent of the magnitude of the phenotypic effects produced by mutation but depended on the total mutation rate, summed over the loci, and the intensity of selection. On the other hand, in the infinitely many allele model of Kimura<sup>14</sup>, the equilibrium genetic variance depended on the underlying biological parameters and the distribution of allelic effects was found to be Gaussian approximately.

The investigations of Turelli<sup>27-29</sup> were based on an alternative approximation for the continuum-of-alleles model of Crow and Kimura<sup>5</sup> on the empirically motivated assumption that the effects of new mutations at a locus are usually much greater than the existing genetic variance at the locus. This led to the prediction of the same equilibrium genetic variance as those for diallelic loci, predicated by Latter<sup>22</sup> and Bulmer<sup>2,3</sup>. Subsequently, Nagylaki<sup>24</sup>, Gillespie<sup>11</sup>, Gimelfarb<sup>12</sup> and Barton and Turelli<sup>1</sup> studied similar problems.

In these studies (with the exception of Latter<sup>23</sup>, Lande<sup>17</sup>, Gillespie<sup>11</sup>, and Barton and Turelli<sup>1</sup>), the primary concern was to explain the magnitude of genetic variance within a population and little consideration was given to the genetic differentiation between populations or species. Moreover, none of these workers adopted a more realistic model of mutation involving discrete change of state. Chakraborty and Nei<sup>4</sup> however, developed a new mutation model called "discrete allelic effect" to examine the extent of



genetic variation of a quantitative trait within a population as well as the same between two populations during the process of their genetic differentiation. But their study considered the forces of mutation and random drift only and selection was ignored. Narain and Chakraborty<sup>25</sup> therefore, studied the evolutionary changes of genetic variance within and between populations for quantitative characters determined by a few loci with major effects by using the discrete-time, discrete allelic-state model with mutation and selection in an infinitely large population. The selection was of the optimum type so that we could examine the change in variance under two situations : (a) a population that evolves from monomorphism (at an optimum phenotype), and (b) when an equilibrium population shifts to a new environment where the optimum phenotype is shifted by a few units of the phenotypic scale. The transient behaviour of the approach to equilibrium was studied in terms of changes in the means as well as the variances. The change in the interpopulational variance was examined when the equilibrium population splitted into two, where one of them moved to a new optimum environment.

In this paper, we critically review some of the models referred to above for studying the dynamics of quantitative characters, both in terms of the means as well as the variances, under mutation-selection balance. In addition, we also study the genetic differentiation between populations or species in respect of quantitative characters.

### ALTERNATIVE APPROACHES

The evolution of quantitative characters can be studied in either of the two ways. One way would be to define and study the underlying models at the level of phenotype avoiding any reference to gene frequencies. The Gaussian phenotypic analyses popularized by Lande<sup>18,20</sup> adopt this approach, on the assumption that genetic variances and covariances are known. Such an approach however, cannot predict the evolutionary dynamics of variance. The other way could be to define and study a genetic model assuming that a complete genetic analysis of the traits is possible. In the latter case, one has to start from the simplest situation of a single locus with two alleles and build over it the more complex systems of infinitely many alleles at a locus, several loci, linkage and epistatic effects etc. The results obtained from the simpler situations of one or two loci give an insight to the problem, particularly for characters such as skin pigmentation in man which is believed to be controlled by a few loci say 5 to 6 with major effect. The rate of evolution, for such characters, is fairly rapid and

modelling with few loci would be realistic. This latter approach would be mostly adopted in this paper.

### RELATIONSHIP BETWEEN PHENOTYPIC AND GENOTYPIC SELECTION

The genetic properties of a given population are determined by gene and genotypic frequencies and need to be connected to quantitative differences noticed in a metric character at the phenotypic level. We consider a population of individuals with  $k$  genotypes  $G_1, G_2, \dots, G_k$  with frequencies  $f_1, f_2, \dots, f_k$  and record the phenotypic measurements on each of the individual for the given character. Two individuals with the same genotype  $G_i$  may then differ slightly in their measurements due to random effects ascribed to environmental differences. To account for it, we take, for the genotype  $G_i$ , a phenotypic distribution  $F(X_i/G_i)$  where  $X_i$  is the random variable giving measurement on the individuals with genotype  $G_i$ . The mean phenotype for this distribution is then

$$g_i = \int F_i f(X_i/G_i) dX_i \quad (1)$$

We thus have a sequence of  $k$  phenotypic distributions with means  $g_1, g_2, \dots, g_k$  corresponding to  $k$  genotypes but with a common environmental variance  $\sigma_E^2$ . The average and variance of the phenotypic values of the character in the population are then,

$$M = \sum_{i=1}^k f_i g_i$$

$$\sigma_P^2 = \sum_{i=1}^k f_i (g_i - \bar{g})^2 + \sigma_E^2 = \sigma_G^2 + \sigma_E^2 \quad \dots(2)$$

the latter expression indicating that  $\sigma_P^2$  is the sum of the between-genotype variance ( $\sigma_G^2$ ) and the average within-genotypic variance ( $\sigma_E^2$ ) due entirely to environmental effects.

The change of the population mean resulting from the selection is brought about through the changes in the gene frequencies at the loci which influence the character under selection. But since the effects of the loci cannot be individually followed, the changes in the gene frequencies cannot, in practice, be ascertained unless we have some means of translating the phenotypic changes into genetic changes and vice-versa. The selection on the basis of the character  $X$  with a certain intensity induces selection among alleles at individual loci controlling the character. This is expressed in terms of a selection function  $W(X)$  defined as relative selective values of an

individual with measurement  $X$ . We can then use the phenotypic distributions together with  $W(X)$  to determine the relative selective values  $W_i$  conferred on the corresponding genotypes. This is simply the average fitness of individuals with a given genotype. These fitness values can be used, along with gene frequencies, to determine the gene frequency in the next generation. This gene frequency information is then used to describe the mean and variance of the character in the next generation. The difference in the mean values between two successive generations gives the response to natural selection.

We follow the method of Kimura and Crow<sup>15</sup> to establish a general relation between the selection made at the overall phenotypic level and the consequent selection induced at the genotypic level at individual loci. Let  $X_{op}$  be the optimum phenotypic value with maximum fitness. Consider a given locus with two alleles  $A$  and  $a$  segregating in the population with respective frequencies  $(1 - q)$  and  $q$ . Assuming a random mating diploid population, let the average phenotypic values of  $AA$ ,  $Aa$  and  $aa$  individuals be respectively  $X_{11}$ ,  $X_{12}$  and  $X_{22}$ . Measuring the character in units of phenotypic standard deviation  $\sigma_p$ , we denote the density (before selection) and fitness functions by  $f(x)$  and  $w(x)$  respectively. If we take  $X_{op}$  as the origin,  $x = (X - X_{op})/\sigma_p$

Let  $m = (M - X_{op})/\sigma_p$  and  $a_{ij} = (X_{ij} - M)/\sigma_p$  where  $a_{ij}$  is the deviation of the average phenotypic value  $X_{ij}$  from the population mean in  $\sigma_p$  units. Let  $w_{ij}$  be the relative fitness of the genotype with value  $X_{ij}$ , then

$$w_{ij} = \int_{-\infty}^{\infty} w(x) f(x - a_{ij}) dx \quad .(3)$$

This is because the distribution of  $x$  in the sub-population of genotype with value  $X_{ij}$  is shifted by  $a_{ij}$  and the density function in this sub-population before selection is  $f(x - a_{ij})$ . This is a good approximation if the locus under consideration is contributing only a small portion of the total variance. Expanding  $f(x - a_{ij})$  in Taylor series about  $a_{ij} = 0$  and integrating, we get

$$w_{ij} = \beta_0 - a_{ij} \beta_1 + \frac{a_{ij}^2}{2} \beta_2 \quad \dots(4)$$

where we neglect third order terms in  $a_{ij}$ , treating it to be small, and

$$\begin{aligned}\beta_0 &= \int_{-\infty}^{\infty} w(x) f(x) dx \\ \beta_1 &= \int_{-\infty}^{\infty} w(x) f'(x) dx \\ \beta_2 &= \int_{-\infty}^{\infty} w(x) f''(x) dx\end{aligned}\quad \dots(5)$$

If the effect of substituting  $a$  for  $A$  is  $\alpha$ , the average genotypic values of  $AA$ ,  $Aa$  and  $aa$  can be expressed as  $a_{11} = -2q\alpha$ ,  $a_{12} = (1 - 2q)\alpha$  and  $a_{22} = 2(1 - q)\alpha$  respectively. Then

$$\begin{aligned}w_{11} &= \beta_0 - a_{11}\beta_1 + \frac{a_{11}^2}{2}\beta_2 = \beta_0 + 2q\alpha\beta_1 + 2q^2\alpha^2\beta_2 \\ w_{12} &= \beta_0 - a_{12}\beta_1 + \frac{a_{12}^2}{2}\beta_2 = \beta_0 - (1 - 2q)\alpha\beta_1 + \left(\frac{1 - 2q}{2}\right)^2\alpha^2\beta_2 \\ w_{22} &= \beta_0 - a_{22}\beta_1 + \frac{a_{22}^2}{2}\beta_2 = \beta_0 - 2(1 - q)\alpha\beta_1 + 2(1 - q)^2\alpha^2\beta_2\end{aligned}\quad \dots(6)$$

The mean fitness of the whole population is

$$\begin{aligned}\bar{w} &= (1 - q)^2 w_{11} + 2q(1 - q) w_{12} + q^2 w_{22} \\ &= \beta_0 + q(1 - q)\alpha^2\beta_2\end{aligned}\quad \dots(7)$$

The gene frequency of  $a$  in the next generation is then

$$q' = (q^2 w_{22} + q(1 - q) w_{12}) / \bar{w}\quad \dots(8)$$

so that the change in gene frequency due to natural selection is

$$\begin{aligned}\Delta q &= q' - q \\ &= (q^2 w_{22} + q(1 - q) w_{12} - q\bar{w}) / \bar{w} \\ &= q(1 - q) [-\alpha\beta_1 + \alpha^2\beta_2 (\frac{1}{2} - q)] / \beta_0 = sq(1 - q)\end{aligned}\quad \dots(9)$$

neglecting terms involving  $\alpha^3$  and higher powers. This gives

$$s = -\alpha (\beta_1/\beta_0) + \alpha^2 \beta_2 (\frac{1}{2} - q)/\beta_0 \quad \dots(10)$$

### MUTATION-SELECTION BALANCE AT A DIALLELIC LOCUS

Stabilising selection reduces genetic variability, but in natural populations this is countered in each generation by fresh variability generated by mutation. We have therefore, to determine how much genetic variability is likely to be maintained by the balance between these two forces. Following Bulmer<sup>3</sup>, we allow a mutation rate  $\nu$  from  $A$  to  $a$  as well as an equal mutation rate from  $a$  to  $A$ . The expression for change in gene frequency given by (9) then becomes

$$\Delta q = q(1 - q) [-\alpha(\beta_1/\beta_0) + \frac{1}{2} \alpha^2 (\beta_2/\beta_0) (1 - 2q)] + \nu(1 - 2q) \quad \dots(11)$$

Let the fitness function be

$$w(x) = \exp \left[ -\frac{1}{2} x^2 / \sigma_w^2 \right] \quad \dots(12)$$

and suppose the mean is zero when  $q = \frac{1}{2}$  at all the loci controlling the trait. At equilibrium with  $q = \hat{q}$  the  $q$ 's must be symmetrical about  $1/2$ .

Then

$$\begin{aligned} (\beta_1/\beta_0) &= 0 \\ (\beta_2/\beta_0) &\cong -1/(\sigma_p^2 + \sigma_w^2) \end{aligned} \quad \dots(13)$$

This gives

$$(1 - 2\hat{q}) \left[ -\frac{1}{2} \alpha^2 q(1 - q) / (\sigma_p^2 + \sigma_w^2) + \nu \right] = 0 \quad \dots(14)$$

Thus  $\hat{q}$  is either  $1/2$  or it satisfies the quadratic equation

$$\alpha^2 \hat{q}(1 - \hat{q}) = 2\nu(\sigma_p^2 + \sigma_w^2) \quad \dots(15)$$

Suppose now that there is an even number of loci ( $2k$ ) and that (15) holds at all loci, half of the gene frequencies being at the smaller and half at the larger root. Since

$\sigma_g^2 = 2k\alpha^2 \hat{q}(1 - \hat{q})$ , we have

$$\begin{aligned} \hat{\sigma}_g^2 &= 4k\nu(\sigma_E^2 + \sigma_w^2)/(1 - 8k\nu) \\ &\cong 4k\nu(\sigma_E^2 + \sigma_w^2) \\ &= 2k\nu/s \end{aligned} \quad \dots(16)$$

As an illustration, suppose  $\sigma_w^2 = 10 \sigma_E^2$  (weak selection) and that  $4kv = 0.01$  implying 250 loci each with a mutation rate of  $10^{-5}$ , the expressed heritability under these rather favourable situations is only 0.10. The important point to note is that the equilibrium genetic variance is independent of the allelic effect. It depends only on the mutation rate and the intensity of selection.

### BALANCE IN AN INFINITELY MANY ALLELE MODEL

The mutation model considered by Kimura<sup>14</sup> allowed for an infinitely many alleles at each locus. In this model, when a mutation occurs, it changes the value of the contribution of that allele by a small amount  $\xi$  from  $x$  to  $(x + \xi)$  where  $\xi$  is a random variable with density function  $f(\xi)$  which is symmetrical about zero with variance  $\sigma^2$ . A continuous time model is further assumed so that the fitness of an individual with value  $x$  is  $-\frac{1}{2} x^2 / \sigma_w^2$  where the optimum value is set at zero. This is equivalent to the usual model of stabilising selection with fitness  $\exp(-\frac{1}{2} x^2 / \sigma_w^2)$  in discrete time.

Let  $p(x, t)$  be the relative frequency of alleles with effect  $x$  at time  $t$ . The rate of change of  $p(x, t)$  due to mutation is

$$\begin{aligned} \frac{\partial p(x, t)}{\partial t} &= -vp(x, t) + v \int_{-\infty}^{\infty} p(x - \xi, t) f(\xi) d\xi \\ &= v\sigma^2 \frac{\partial^2 p(x, t)}{\partial x^2} + o(\xi^2) \end{aligned} \quad \dots(17)$$

where  $v$  is mutation rate per generation,  $p(x - \xi, t)$  is expanded around  $\xi = 0$  by Taylor series and

$$\begin{aligned} \int_{-\infty}^{\infty} \xi f(\xi) d\xi &= 0, \\ \int_{-\infty}^{\infty} \xi^2 f(\xi) d\xi &= \sigma^2, \\ \int_{-\infty}^{\infty} \xi^n f(\xi) d\xi &= 0, n > 3 \end{aligned} \quad \dots(18)$$

The rate of change of  $p(x, t)$  due to selection is

$$\begin{aligned} \frac{\partial p(x, t)}{\partial t} &= \frac{1}{2} p(x, t) \left[ -x^2 + \int_{-\infty}^{\infty} x^2 p(x, t) dx \right] / \sigma_w^2 \\ &\quad - \frac{1}{2} p(x, t) [\sigma_g^2 - x^2] / \sigma_w^2 \end{aligned} \quad \dots(19)$$

where

$$\begin{aligned} \int_{-\infty}^{\infty} xp(x, t) dx &\rightarrow 0 \\ \int_{-\infty}^{\infty} x^2 p(x, t) dx &\rightarrow \hat{\sigma}_g^2 \\ p(x, t) &\rightarrow p(x) \end{aligned}$$

as  $t \rightarrow \infty$  in the steady state. At equilibrium, therefore, we have,  $\partial p(x, t) / \partial t = 0$  and

$$\frac{1}{2} v \sigma^2 \frac{d^2 p(x)}{dx^2} - \frac{1}{2} P(x) (x - \hat{\sigma}_g^2) / \sigma_w^2 = 0 \quad \dots(20)$$

This ordinary differential equation is satisfied when  $p(x)$  is the density function of a normal distribution with mean zero and variance given by

$$\hat{\sigma}_g^2 = \sqrt{v \sigma^2 \sigma_w^2} \quad \dots(21)$$

Hence in the steady state, the balance between mutation of an infinitely many allele type and stabilising selection, in continuous time, produces a Gaussian distribution of allelic effects with zero mean and variance given by (21). The genetic value of the individual is therefore, distributed normally around zero mean with variance  $2\hat{\sigma}_g^2$ .

## BALANCE WITH A STEP-WISE DISCRETE MUTATION MODEL

We consider a quantitative character controlled by  $n$  loci and assume that at each locus there is an infinite number of possible allelic states. We assume that the phenotypic effect of the alleles are discrete as shown in Fig. 1.

In this figure,  $A_i$  represents an allele occupying state  $i$  (any integer number from  $-\infty$  to  $\infty$ ) and having an allelic effect of  $ia$ . We assume that all

allelic effects are additive with no dominance and no epistasis and that once  $A_i$  mutates, it changes to allelic state  $A_{i+r}$  with probability.

$$\alpha_r = \alpha_{-r} = \binom{2m}{m-r} \left(\frac{1}{2}\right)^{2m} \text{ for } 0 \leq r \leq m \quad \dots(22)$$

$\alpha_r = 0$ , otherwise

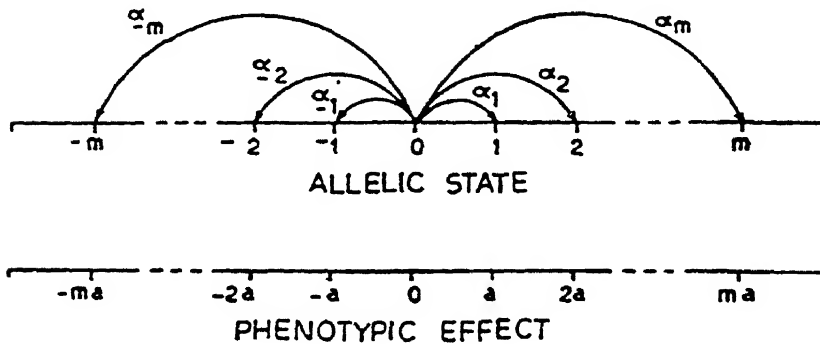


FIG 1 Discrete allelic-state model used in this paper. In this model allele  $A_i$  mutates to  $A_{i+r}$  with probability  $\alpha_r (= \alpha_{-r})$ . Allele  $A_i$  has a phenotypic effect of  $a_i$

where  $m$  is the possible number of mutational steps. The distribution is a shifted Binomial with mean zero and variance  $m/2$ . If  $v$  denotes the mutation rate, the absolute probability of such a mutation would be  $v\alpha_r$ . Thus, an allele that mutates but has the same allelic effect as that of the original allele with probability  $v\alpha_0$  so that in the conventional definition, the real mutation rate  $v$ , would be given by  $v' = (1 - \alpha_0)v$ . The per generation increment of the variance of allelic effect by mutation is then  $VMA^2/2$ .

The selection operates on the total phenotypic value  $x$  and it is assumed to follow a normal distribution with mean  $\mu$  and variance  $\sigma_p^2$ . With no genotype-environment interaction, the distribution of  $x$ , among individuals of type  $A_iA_j$  with genotypic value  $a(i+j)$  follows a normal distribution with mean  $a(i+j)$  and the environmental variance  $\sigma_e^2$  with density function

$$f(x/a(i+j)) = \frac{1}{\sigma_e \sqrt{2\pi}} \exp \left[ -\frac{\{x - a(i+j)\}^2}{2\sigma_e^2} \right] \quad \dots(23)$$

The fitness function for the character value  $x$  is assumed to be Gaussian type as



$$w(x) = w_{\max} \exp \left[ -\frac{(x - x_{\text{opt}})^2}{2\sigma_w^2} \right] \quad \dots(24)$$

where the character assumes the optimum fitness  $w_{\max}$  at  $x = x_{\text{opt}}$  and  $\sigma_w$  is the width of the function indicating the rate at which fitness declines with deviation of  $x$  from the optimum. Taking  $w_{\max} = 1$ , the mean fitness of the individuals with genotype  $A_i A_j$  would be

$$\begin{aligned} w_{ij} &= \int w(x) f(x/a(i+j)) dx \\ &= \sigma_w \sqrt{2s} \exp [-s\{a(i+j) - x_{\text{opt}}\}^2] \end{aligned}$$

where  $s = 1/2 (\sigma_w^2 + \sigma_e^2)$  indicates the strength of the selection at the group level. A large  $\sigma_w$  means weak selection of the stabilising type.

If  $x_i(t)$  denotes the frequency of allele  $A_i$  in generation  $t$  with allelic effect  $ai$ , an individual of genotype  $A_i A_j$  will have a mean reproductive fitness  $w_{ij} = \exp [-sa^2(i+j)^2]$  and thus, the change in gene frequency of  $A_i$  from generation  $t$  to  $t+1$  is given by

$$\begin{aligned} \bar{\omega}_A(t) x_i(t+1) &= (1 - v + v\alpha_0) \sum_j x_i(t) x_j(t) \exp [-sa^2(i+j)^2] \\ &+ v \sum_{r=1}^m \alpha_r \left( \sum_j x_j(t) [x_{i+r}(t) \exp \{-sa^2(i+i+r)^2\} \right. \\ &\left. + x_{i-r}(t) \exp \{-sa^2(i+j-r)^2\}] \right) \end{aligned} \quad \dots(26)$$

where  $\bar{\omega}_A(t)$  is the mean fitness of individuals at the locus in the  $t$ th generation so adjusted as to make  $\sum_j x_j(t) = 1$ .

In general, this recurrence relationship does not yield any explicit solution. However, for  $m = 1$  it is possible to derive the equilibrium allele frequency profile by neglecting powers of  $v$  and  $s$  when the mean fitness  $\bar{\omega}_A(t)$  is approximated as

$$\bar{\omega}_A(t) \cong 1 - s \sigma_{g_A}^2(t) \quad \dots(27)$$

where  $\sigma_{g_A}^2(t)$  is the total genotypic variance contributed by this locus at time  $t$  given by

$$\sigma_{g_A}^2(t) = a^2 \sum_i \sum_j x_i(t) x_j(t) (i+j)^2 \quad \dots(28)$$

The recurrence relation reduces to

$$x_i(t) = \frac{\nu}{2} [x_{i+1}(t-1) + x_{i-1}(t-1)] + x_i(t-1) [1 - \nu - s \times \{a^2 i^2 - \sigma_{gA}^2(t-1)/2\}] \quad \dots(29)$$

giving

$$\Delta x_i(t) = -\nu \left[ x_i(t) - \frac{x_{i+1}(t) + x_{i-1}(t)}{2} \right] + s \left[ \frac{\sigma_{gA}^2(t)}{2} - a^2 i^2 \right] \quad \dots(30)$$

When the population reaches equilibrium under the opposing pressures of mutation and selection, we have  $\Delta x_i = 0$ . This gives

$$\begin{aligned} \hat{x}_1 &= (1 - SG) \hat{x}_0 \\ \hat{x}_{i+1} - 2[1 - S(G - i^2)] \hat{x}_i + \hat{x}_{i-1} &= 0, i \geq 1 \\ \hat{x}_i &= \hat{x}_{-i}, \end{aligned} \quad \dots(31)$$

The frequent allele frequencies are obtained as

$$\begin{aligned} \hat{x}_1 &= \hat{x}_0 (1 - SG) \\ \hat{x}_2 &= \hat{x}_0 [1 - 2SG + 2S(1-G)(1-SG)] \end{aligned} \quad \dots(32)$$

and

$$\begin{aligned} \hat{x}_3 &= \hat{x}_0 [1 - 3SG + 6S(1-G)(1-SG) \\ &\quad + 2S\{3 - SG(6-G) + 2S(1-G) \\ &\quad \times (4-G)(1-SG)\}] \end{aligned}$$

$$\text{where } S = sa^2/\nu \text{ and } G = \frac{1}{2} \sum_i \sum_j \hat{x}_i \hat{x}_j (i+j)^2 = \sigma_{gA}^2 / 2a^2$$

In the general case of  $m$ -step mutational changes, the moments of the allelic effects as well as those of genotypic effects can be obtained analytically under optimum selection. Denoting the  $k$ th moment of the distribution of allelic effects at a locus in the  $t$ th generation by

$$M_k(t) = \sum_{i=-\infty}^{\infty} a^k i^k x_i(t) \text{ and noting that for all } i, x_i(i) = x_{-i}(t) \text{ at each}$$

generation (since the optimum genotype is at origin), the recurrence relationship for the even order moments is given by

$$\begin{aligned} M_{2k}(t+1) &= [1 - \nu + \nu \left( \frac{2m}{m} \right) \left( \frac{1}{2} \right)^{2m}] M_{2k}(t) \\ &\quad + s[M_2(t) M_{2k}(t) - M_{2k+2}(t)] \end{aligned}$$

$$+ 2v \sum_{i=0}^k \binom{2k}{2i} M_{2i}(t) \sum_{r=1}^m \binom{2m}{m-r} \left(\frac{1}{2}\right)^{2m} (ar)^{2k-2i} \dots(33)$$

so that the change of variance of allelic effects at generation  $t$ ,  $\Delta M_2(t)$  is given by

$$\Delta M_2(t) = \frac{mva^2}{2} + s [M_2^2(t) - M_4(t)] \dots(34)$$

At equilibrium, therefore, the fourth moment and the variance of allelic effects are related by

$$\widehat{M}_4 = \widehat{M}_2^2 + \frac{mva^2}{2s} \dots(35)$$

Now, if  $\mu_r(t)$  denotes the  $r$ th order moment of the genotypic effects at the locus (i.e.,  $\mu_2(t) = \sum_i \sum_j a^r (i+j)^r x_i(t) x_j(t)$ ), the variance and the fourth moment of the genotypic effect at a locus are related with those of allelic effects by

$$\mu_2(t) = 2M_2(t) \text{ and } \mu_4(t) = 2M_4(t) + 6M_2^2(t)$$

We thus obtain the change of variance of genotypic values at a locus in generation  $t$ ,  $\Delta \mu_2(t)$ , as

$$\Delta \mu_2(t) = mva^2 - s [\mu_4(t) - 2\mu_2^2(t)] \dots(36)$$

If the equilibrium distribution of genotypic values is normal (i. e.,  $\widehat{\mu}_4 = 3\widehat{\mu}_2^2$ ) the genotypic variance at a locus in the equilibrium population is given by  $\widehat{\sigma}_{gA}^2 = \widehat{\mu}_2 = \sqrt{mva^2/s} = \sqrt{\Delta \sigma_m^2/s}$  where  $\Delta \sigma_m^2 = mva^2$  is the effect of mutational change on the genotypic value at a locus. This result is identical to that of Kimura<sup>14</sup> even though his model assumes a continuous distribution of allelic effects.

Under the discrete allelic effect model, the departure from normality can be determined from the value of  $\widehat{\beta}_2 = \widehat{\mu}_4/\widehat{\mu}_2^2$  at equilibrium. We thus obtain

$$\widehat{\beta}_2 = 2 + \Delta \sigma_m^2/s \widehat{\sigma}_{gA}^4$$

and hence the departure from normality as measured by  $\gamma = \widehat{\beta}_2 - 3$  is given by

$$\gamma = (\Delta\sigma_m^2 - s\hat{\sigma}_{gA}^4)/s\hat{\sigma}_{gA}^4$$

### GENETIC DIFFERENTIATION BETWEEN POPULATIONS OR SPECIES

Let us now consider the case where the population which is at equilibrium initially with mean phenotype at origin and variance  $\sigma_p^2$  now shifts to a new environment where the optimum phenotype is  $d\sigma_p$  units away from the mean. Obviously, under the effect of mutation and selection the equilibrium status of genotypic distribution will be immediately disturbed and gradually the distribution will shift towards the new optimum. To analyze the nature and rate of change of variability we must again consider the recurrence relationship of gene frequency changes. In this new environment, the fitness,  $w_{ij}$ , of an individual of genotype  $A_1A_j$  is given by

$$w_{ij} = \exp[-s\{a(i+j) - d\sigma_p\}^2].$$

We then get, following similar derivations,

$$\mu'_1(t+1) = \mu'_1(t) - s[\mu_3(t) + 2\mu_2(t)\{\mu'_1(t) - d\sigma_p\}] \quad \dots(38)$$

where  $\mu'_r(t)$  is the  $r$ th order moment (about origin) of the genotypic values in the  $t$ th generation. Clearly, at equilibrium we thus have  $\mu'_1 = d\sigma_p$ , the optimum genotypic value and  $\mu_3 = 0$ . The equilibrium distribution being symmetric, the change in genotypic variance is, similarly, given by

$$\Delta\sigma_{gA}^2(t) = mva^2 + s[2\sigma_{gA}^4(t) + \mu_3(t)\{(d\sigma_p) - 2\mu'_1(t)\} - \mu_4(t)] \quad \dots(39)$$

where  $\mu_3(t)$ ,  $\mu_4(t)$  represent the third and fourth moments (about mean) of genotypic values at the  $t$ th generation. At equilibrium, since  $\mu_3$  is zero, we again have the same steady state genotypic variance given by

$$\sigma_{gA}^2 = \sqrt{\frac{1}{2}\left\{\hat{\mu}_4 - \frac{mva^2}{s}\right\}} \quad \dots(40)$$

Thus, even if the optimum is shifted by a certain s. d. away from the original mean, as long as the intensity of selection ( $s$ ) remains the same, the genotypic variance eventually returns to its original value although the genotypic distribution becomes now centered around the new optimum genotypic value.

At the transitory stage however, it is difficult to assert analytically how the variance is altered. However, as we shall see in our numerical computations, at a transitory stage the genotypic variance first increases and eventually returns to its original equilibrium value.

It is apparent from the theoretical development given in the previous section that a quantitative study of genetic differentiation for metric traits between populations or species cannot be made analytically if optimal selection with stepwise mutation in an infinite population is envisaged. A computer was therefore used to compute numerically the various quantities of interest by resorting to exact recurrence relations already discussed in the previous sections. The mean and variability of the character both within as well as between populations were studied in the transient stage and at equilibrium. To start with, the initial population was considered as monomorphic at optimum and the behaviour of within population variance studied over time. After reaching the equilibrium, the optimum was shifted to a few units on the right and the transient behaviour of the mean and variability (within as well as between) was studied.

Since the mean is at the optimum and the optimum is set zero, the population mean remains at zero unless there is a shift of the optimum genotype. However, the variance increases slowly from zero and attains, at equilibrium, a value determined solely by  $(v/s)$  as already obtained algebraically. This behaviour of within population variance as a function of time for different values of  $s$  between 0.004 to 0.040 for  $m = 1$  and  $v = 0.0005$  is presented in Table I.

For intense selection, the variances during the transient stage as well as at equilibrium are lower as otherwise expected. Mutation creates variability while selection eliminates it so that for intense selection, its role is dominant. Also, the approach to equilibrium is quicker for more intense selection as revealed by the results of computer simulations which are not presented here.

**Table I**

*Within population variance at equilibrium when initial population is genetically homogeneous with only one type of individuals which is optimum phenotype for some selected values of selection coefficients ( $s$ )*  
*The mutation rate,  $v$ , is taken as 0.0005 and  $m = 1$*

| $s$   | $\hat{\sigma}_{gA}^2$ |
|-------|-----------------------|
| 0.004 | $1.17 \times 10^{-1}$ |
| 0.008 | $6.03 \times 10^{-2}$ |
| 0.010 | $4.85 \times 10^{-2}$ |
| 0.020 | $2.44 \times 10^{-2}$ |
| 0.040 | $1.22 \times 10^{-2}$ |

When we shift the optimum to four or six standard deviations away from the mean on the right and study the transient behaviour of the process as it approaches the same equilibrium, we notice some interesting results. In Table II, we present these results in terms of mean, variance, skewness and kurtosis of genotypic values in different generations for a quantitative character under centripetal selection when the new optimum is at six standard deviations away from the original optimum. We take  $v = 0.001$ ,  $s = 2v$ , and  $m = 5$ .

**Table II**

*Mean ( $\bar{X}_t$ ), variance ( $V_t$ ), skewness ( $\gamma_{3t}$ ) and kurtosis ( $\gamma_{4t}$ ) of genotypic values in different generations for a quantitative trait under centripetal selection ( $v = 0.001$ ,  $s = 2v$  and  $m = 5$ ) with new optimum 6 standard deviations away from original optimum*

|                            | <i>t</i> |       |       |       |         |          |
|----------------------------|----------|-------|-------|-------|---------|----------|
|                            | 0        | 50    | 100   | 500   | 1000    | $\infty$ |
| Mean ( $\bar{X}_t$ )       | 0        | 0.378 | 0.930 | 3.443 | 3.811   | 4.000    |
| Variance ( $V_t$ )         | 0.456    | 0.728 | 1.296 | 0.815 | 0.618   | 0.456    |
| Skewness ( $\gamma_{3t}$ ) | 0        | 0.997 | 0.684 | 0.119 | - 0.001 | 0        |
| Kurtosis ( $\gamma_{4t}$ ) | 2.741    | 1.886 | 0.433 | 0.764 | 1.563   | 2.741    |

N. B.: Kurtosis ( $\gamma_{4t}$ ) is measured by  $\gamma_{4t} = (\mu_{4t}/V_t^2)^{-3}$

The mean increases from zero to four at equilibrium but the increase is more rapid for a higher value of  $m$ . The variance on the other hand increases, attains a transitory maximum and decreases back to the original steady state value which is related to the fourth moment of the genotypic values as given by (40). When  $s$  is very large compared to  $V$ , more than one transitory maxima are produced. In Fig. 2, we present such a behaviour of intra-population variance for  $s = 0.002$ ,  $0.004$  and  $0.008$  when  $V = 0.001$ ,  $m = 1$  and when the new optimum is at six standard deviations away from the original optimum at the origin.

For a higher value of  $m$ , the variance attains a considerably higher peak as well as somewhat earlier than when  $m = 1$ . The most interesting feature is regarding the skewness of the distribution of genotypic values. Initially, this distribution is symmetrical but as we advance in time, its symmetry is disturbed. It gets skewed initially and then slowly the skewness decreases, changes sign and finally the distribution becomes again symmetrical at equilibrium. The kurtosis of the distribution also behaves in a similar fashion. Starting from a value very near to three

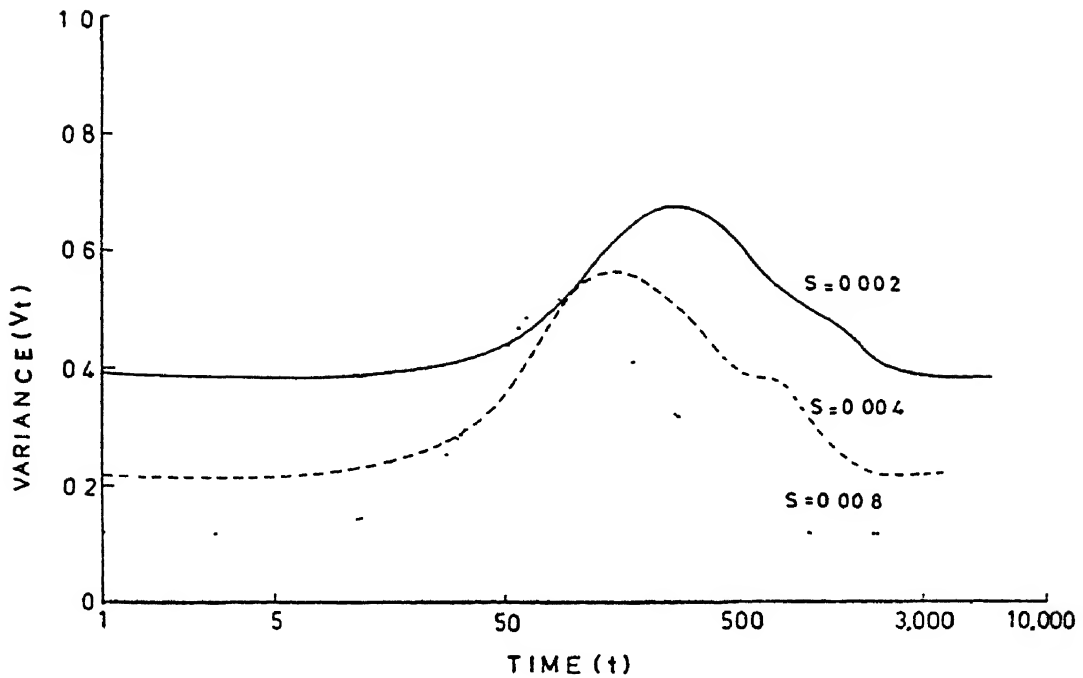


FIG 2 Intra-population variance ( $V_t$ ) under the joint effects of mutation and centripetal selection as function of time  $t$ , with  $m = 1$ ,  $v = 0.001$ ,  $s = 0.002$ ,  $0.004$  and  $0.008$  when the new optimum is at 6 standard deviations away from the original optimum at the origin. Initial population is at steady-state

initially, it declines to a value less than half but increases thereafter and restores the initial value at the equilibrium. Compared to  $m = 1$ , the distribution for  $m = 5$  becomes more leptokurtic. A typical allelic distribution corresponding to  $v = 0.001$ ,  $s = 2v$  and  $m = 5$ , depicting these features, is shown in Fig. 3.

Of special importance in these studies is the genetic differentiation between populations built up over a period of time when in one population the same optimum holds but in the other it has shifted a certain distance away from the mean. Chakraborty and Nei<sup>4</sup> used the ratio of between population ( $B_t$ ) within population ( $V_t$ ) variability as an index for determining the evolutionary forces under which the character changes over time. Their studies involving drift only showed that this ratio ( $B_t/V_t$ ) increases linearly with time. With centripetal selection in an infinitely large population we find that this behaviour changes considerably and it is no longer a monotone function of time.

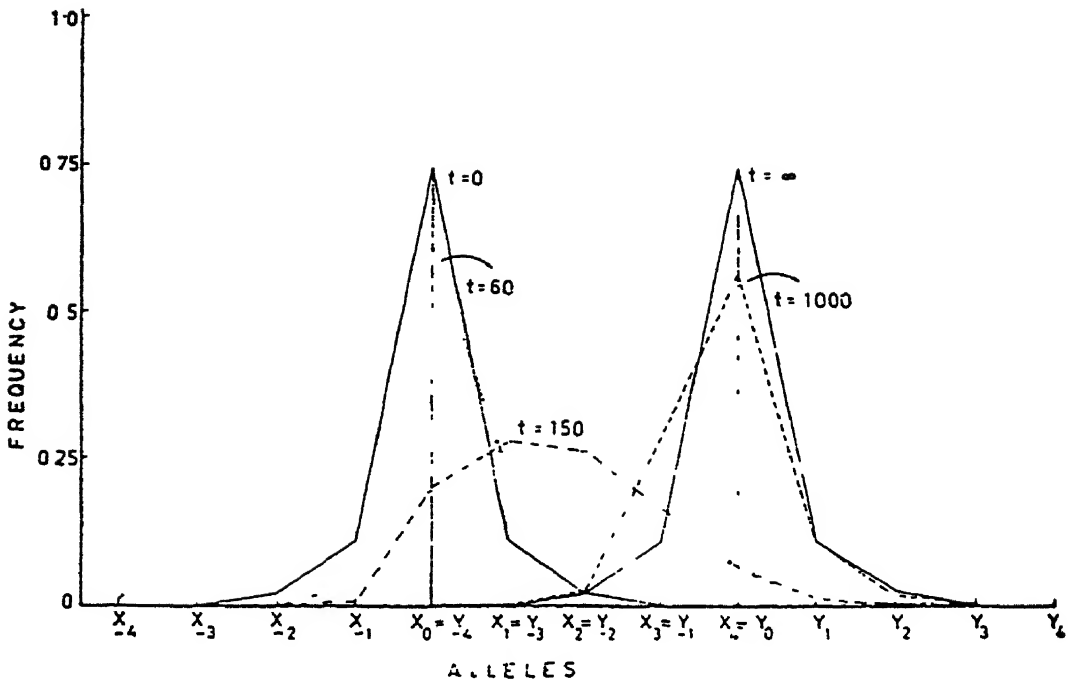


FIG 3 Allele frequency distribution under joint effects of mutation and centripetal selection at different generations ( $t = 0, 60, 150, 1000$  and  $\infty$ ) with  $\nu = 0.001$ ,  $s = 2\nu$  and  $m = 5$  when the new optimum is 6 standard deviations away from the original optimum at the origin. Initial population is at steady-state.

We have already discussed the behaviour of  $V_t$  which increases slowly from initial equilibrium value, attains a maximum and then decreases back to the same value at equilibrium. But when we consider between population variability,  $B_t$ , it is found that it increases slowly initially and then almost linearly until it approaches a plateau at equilibrium. The ratio  $(B_t/V_t)$  almost mimics the behaviour of  $B_t$  at least in the initial stages but it attains a much higher value. This is obvious because  $V_t$  decreases while  $B_t$  increases as equilibrium is reached. In the initial transient stage, however,  $B_t$  and the ratio are almost the same because  $V_t$  has been increasing and reaching a maximum. After this stage, at which maximum  $V_t$  occurs, the quantities  $B_t$  and  $(B_t/V_t)$  diverge apart, increasing with time by different magnitudes. To illustrate the qualitative nature of the changes in  $V_t$ ,  $B_t$  and  $(B_t/V_t)$ , under joint effects of mutation and centripetal selection, Fig. 4 presents the numerical results for  $\nu = 0.001$ ,  $s = 2\nu$  and  $m = 5$  when the new optimum is taken to be approximately six standard deviations away from the optimum in the other population.



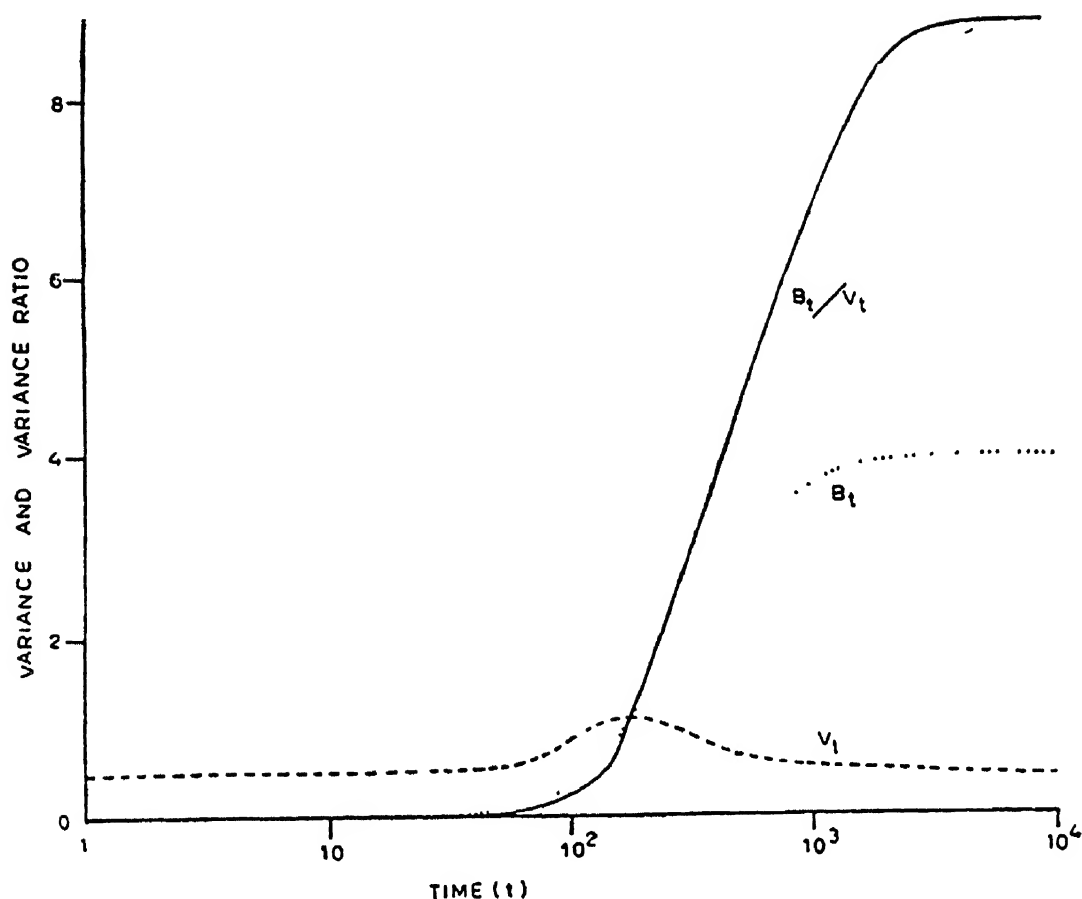


FIG 4 Intra- ( $V_t$ ) and inter-population ( $B_t$ ) variance components under the joint effects of mutation and centripetal selection as a function of time of divergence of two populations one of which has optimum genotype six standard deviations away from the other (which remains at a steady state by mutation-selection balance). Time ( $t$ ) is measured in units of generation. Initial population is at steady state. The parameter values are  $v = 0.001$ ,  $s = 2v$  and  $m = 5$ .

It is seen from this figure that the rate of approach to equilibrium variance within population (0.456) is faster as compared to that of the mean genotypic value (4.0) since the between-population variance ( $B_t$ ) attains its equilibrium value (4.0) at a later time than  $V_t$ . The variance ratio ( $B_t/V_t$ ) reaches a steady state value (8.80) since the process of genetic differentiation stops once the diverging population reaches its steady-state genotypic distribution around its new optimum value. We thus see that ( $B_t/V_t$ ) asymptotes and does not therefore increase linearly with time as in Chakraborty and Nei<sup>4</sup> under mutation-drift balance. The behaviour of ( $B_t/V_t$ ) as a function of time of divergence can therefore be taken as a criterion for determining whether selective forces are operating or not.

## DISCUSSION

Kimura<sup>14</sup> showed, for infinitely large populations, and assuming a continuous time process that the distribution of allelic effects tends to be normal at equilibrium between selection and mutational forces and that the mean and variance of the equilibrium distribution are determined by the amounts of increase in mean and variance of the genotypic value per gene per generation as well as by the intensity of fitness function. The discrete allelic-state model with the assumption of a discrete-time process, considered in this paper, has revealed behaviour similar to those of Kimura<sup>14</sup> as it should, since binomial distribution of allelic effects should tend to normal distribution as we go from discrete to continuous case. On the other hand, the diallelic model of Latter<sup>22</sup>, reanalysed by Bulmer<sup>2,3</sup> who used different approximations but arrived at the same conclusions as those of Latter<sup>22</sup>, indicated different results. Their predictions for equilibrium genetic variance differ qualitatively from those of Kimura-Lande-Fleming-Narain & Chakraborty.

Based on "House-of-cards" approximation of Kingman<sup>16</sup>, Turelli<sup>27</sup> presented a new asymptotic analysis of Kimura's model to show that the qualitatively different predictions about equilibrium genetic variance are not due to the number of alleles assumed per locus. Instead, such different results are attributable to assumptions concerning relative magnitudes of per locus mutation rates, the phenotypic effects of mutation and intensity of selection. He then analysed a model with tri-allelic loci which allows among-locus variation in mutation rates and allelic effects. At the single locus level, such a model is a particular case of the discrete-time, discrete allelic-state mutation model analysed in the present paper and leads to the same conclusions as in Turelli<sup>27</sup> in so far as the equilibrium genetic variance is concerned.

As already mentioned in the Introduction, very few studies on the problem of genetic differentiation between population or species have appeared, particularly for the situation when one of the daughter populations has a shifted optimum. Such cases have biological relevance as for instance in skin pigmentation for a small group of Caucasian race with fair skin who moved out of Central Asia around 3000 years ago and settled in southern parts of America with plenty of sunlight. The optimum phenotype for skin pigmentation must therefore, have shifted by several standard deviations away from the original optimum. This introduces differentiation between populations and it is of interest to study the transient properties of such a

process. This has been done in this paper by studying the distribution of allelic frequency as a function of time. Significant changes in statistical properties of the distribution such as mean, variance, skewness and kurtosis have been noticed. In particular, algebraic expressions for changes in mean and variance reveal interesting results. When we consider optimum at  $d$  standard deviations units away from the origin, the mean change in the genotypic value of the character at the  $t$ -th generation, as given by eqn. (38), depends on  $s$ ,  $d$ ,  $\sigma_p$ ,  $\mu_3(t-1)$ ,  $\mu_2(t-1)$  and  $\mu'_1(t-1)$ . It does not depend on the mutation rate,  $v$ . But change in the genotypic variance per generation, as given by eqn. (39), clearly indicates that it is affected by the mutational component. Initially, the population is in equilibrium and symmetrical. If we further assume that it has normal kurtosis, we have  $\mu'_1(0) = 0$ ,  $\mu_3(0) = 0$ ,  $\mu_4(0) = 3\mu_2^2(0)$ . If we also disregard the increase in variation due to mutation, the expression (39) reduces to the result given in Latter<sup>23</sup>. The genotypic variance is reduced and the reduction depends on the intensity of selection and the heritability.

The transient behaviour of the mean noticed in this investigation can be useful in giving some idea about the divergence time, at least the time by which the mean gets to half of the total change in the mean, i.e.,  $1/2(d\sigma_p)$ , which is known. The computer results show that intense selection speeds up the divergence time but by increasing the number of mutational steps, this time is shortened, though not very appreciably. Thus, for  $v = 0.001$ ,  $s = 0.008$  and  $m = 1$ , this time is of the order of 120 generations. With  $s = 0.002$  and  $m = 1$  and 5, the times are respectively 300 and 170 generations.

The main focus of this study is on the genetic differentiation between populations which gets built up over time when the population, after reaching equilibrium by mutation-selection balance, splits up into two in one of which the same optimum holds but in the other it shifts a few standard deviations away from the mean. The ratio between versus within population variance ( $B_i/V_i$ ) as a function of time then provides with a possible mean of ascertaining the role of adaptive changes under which the character changes over time. With selection, this ratio necessarily changes nonlinearly with time.

## SUMMARY

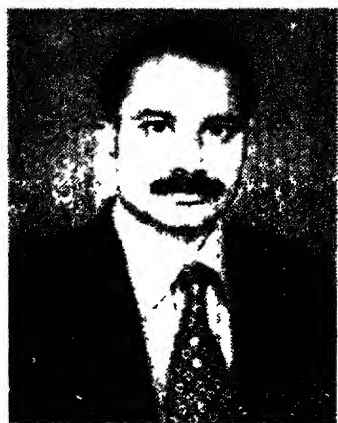
Studies on the maintenance of genetic variability for a quantitative trait due to a balance between stabilising selection and mutation in natural

populations have been reviewed. Several models, both in terms of the dynamics of the means as well as the variances have been discussed. In particular, using a new and more general genetic model called the discrete-allelic state model and assuming discrete-time process, the evolutionary changes of genetic variation of quantitative characters, controlled by a few loci, within and between populations during the process of genetic differentiation of populations or species, are studied under the effects of mutation and centripetal selection in infinitely large populations. While in a finite population and ignoring selection, the rate of change of additive genetic variance depends on mutation and effective population size, traits under optimal selection in infinitely large populations go through the dynamics of a rather complicated form depending on the relative intensities of selection and mutation. When a population, which has reached steady-state by mutation-selection balance, splits into two, in one of which the same optimum genotype holds but in the other the optimum shifts a few standard deviations away from the original optimum, the corresponding daughter population starts differentiating from its sister population by favouring certain class of mutant alleles and discarding others which were originally favoured. During this process of turn over of genes, both the intra-and inter-population variances undergo a complicated change, and the ratio of the latter to the former is a non-linear function of time of divergence. This pattern is qualitatively very different from the case when selection is absent. The intra-population distribution of genotypic values, during this transition, is shown to deviate considerably from normality.

## References

1. N H Barton, and M Turelli, *Genet. Res., Camb* **49** (1987) 157-73
2. M G Bulmer, *Genet Res* **19** (1972), 17-25
3. M G Bulmer, *The Mathematical Theory of Quantitative Genetics* Clarendon Press, Oxford, 1980
4. R Chakraborty, and M Nei, *Genet. Res , Camb.* **39** (1982), 303-14
5. J F Crow, and M Kimura, In *Proc. XI Int. Congr. Genet.* (1964), pp. 495-505
6. C Darwin, *The Origin of Species by Means of Natural Selection* John Murray, London, 1859
7. R A Fisher, *Trans R. Soc. Edinb* **52** (1918), 399-433
8. R A Fisher, *Proc. R. Soc. Edinb.* **42** (1922), 321-41

9. R A Fisher, *The Genetical Theory of Natural Selection*. Oxford University Press
10. W H Fleming, *SIAM J. appl. Math.* **36** (1979), 148-68
11. J H Gillespie, *Genetics* **107** (1984), 321-30
12. A Gimelfarb, *Genetics* **112** (1986), 717-25
13. J B S Haldane, *The Causes of Evolution*, Longmans, London, 1932
14. M Kimura, *Proc Natn. Acad Sci U S A* **54** (1965), 731-736
15. M Kimura, and J F Crow, *Proc Natn. Acad Sci , U S.A.*, **75** (1978), 6168-71
16. J F C Kingman, *J appl Prob* **15** (1978), 1-12
17. R Lande, *Genet Res , Camb* **26** (1975), 221-35
18. R Lande, *Evolution* **30** (1976), 314-34
19. R Lande, *Genetics* **86** (1977), 485-98
20. R Lande, *Evolution* **33** (1979), 402-16
21. R Lande, *Genetics* **94** (1980), 203-15
22. B D H Latter, *Aust J Biol Sci* **13** (1960), 30-35
23. B D H latter, *Genetics* **66** (1970), 165-86
24. T Nagylaki, In *Human Population Genetics* (ed. A Chakravarti). The Pittsburgh Sympo. Hutchinson Ross., Stroudsburgh, Pa., 1984, pp. 275-306
25. P Narain, and R Chakraborty, *Heredity* **59** (1987) 199-212
26. A Robertson, *J Genet* **54** (1956), 236-48
27. M Turelli, *Theor Pop Biol* **25** (1984), 138-193
28. M Turelli, *Genetics* **111** (1985), 165-95.
29. M Turelli, In *Evolutionary Processes and Theory* (ed.: S Karlin and E Nevo), Academic Press, New York, 1986, pp. 607-628
30. S Wright, *Genetics* **16** (1931), 97-159
31. S Wright, *Proc Natn. Acad Sci. U S A.* **23** (1937), 307-20



**Ramamurti Rajaraman** (b. 11 March 1939) did Ph.D. (1963) from University of Cornell, USA. He is Professor at Centre for Theoretical Studies (1976- ), MSIL Chair (1991- ), Indian Institute of Science, Bangalore.

Rajaraman's contributions relate to: three-body theory in nuclear matter; particle physics phenomenology and electroproduction, theory of charged quantum solitons; methods for finding new solitons; first quantization of anomalous gauge theory; proof of absence of topological anomaly, and clarification of fractional charge in polyacetylene. He has

authored 'Introduction to Solitons and Instantons in Quantum Field Theory' (North-Holland, 1982) and a monograph 'Some Non-Perturbative Semi-classical Methods in Quantum Field Theory'.

Rajaraman is Fellow of Indian Academy of Sciences. He is the recipient of Shanti Swarup Bhatnagar Prize (1983); G.P. Chatterjee Memorial Lecture Award (INSA) (1989).

*Ramamurti Rajaraman was elected to the fellowship of the Academy in 1985.*

**Address:** Professor, Centre for Theoretical Studies, Indian Institute of Science, Bangalore 560 012, Tel. (080) 344411 (Ext 2474, 2266); Res: 57/2, 17A Cross, 8th Main, Malleswaram, Bangalore 560 055; Tel: (080) 348505.

## **ELEMENTARITY**

R RAJARAMAN

Although Elementary Particle Physics was sanctified as a separate branch of Physics only about 40 years ago, the pursuit of Elementarity has been on ever since science began—starting from ancient civilisations. Elementarity involves the search for the underlying constituents, the basic "building blocks" out of which all things in the universe are made. This pursuit has been based on the belief, or hope, that just as all the things we see in day-to-day life are made of smaller "parts", which in turn are made of still smaller parts and so on, if we keep pushing this to tinier and tinier scales we will eventually unearth some small number of building blocks or "constituents" out of which are made all the myriad objects that we find in our universe.

This search for the ultimate constituents of all things is not just of curiosity value. To the extent that we are able to discover that everything is made up of a relatively small number of constituents, the relatively few laws which describe the behavior of these relatively few basic constituents would, at least in principle, be responsible for all the behavior of all the billions of different objects we encounter in the world. This is not only conceptually very satisfying, but could also simplify our understanding of the physical world and lead to practical dividends. Indeed in some sense much of modern science and technology emerged from this pursuit of elementarity.

In this talk I will first give a capsule history of how the pursuit of elementarity has progressed so far, and notice a pattern in the way this history has evolved. This pattern will already cast doubts on the whether this search of basic constituents of all matter will ever end, or whether it is doomed to be an endless quest. But the bulk of my talk will be devoted to something much more interesting from the conceptual and intellectual viewpoint. I will try to show you that as the pursuit of elementarity entered the second half of this century, the very notion of elementarity began to be threatened. Fit the level of the elem. particles of today our basic picture of "parts" and the "whole", of smaller constituents making up a larger composite may itself have to be abandoned, or at least seriously modified.

But before I discuss all that, here's a capsule history of elementarity. As I said earlier, the pursuit of elementarity has been a running theme in science ever since science began. Theories in ancient civilisations to the effect that all things are made up of earth, fire, water and air may sound amusing in today's context, but they already represent a search for the ultimate constituents of all matter. However, by the time the modern era of science began a few centuries back, it was clear that the world contained thousands of different objects—various minerals, plants, birds, animals and so on, which did not seem to be "made of" just earth, fire, air, and water in any objective, experimentally verifiable sense. By then science had developed these standards, instead of just abstract speculation, as the criteria for acceptability of any scientific theory. So the theme of elementarity was again pursued, by chemists, long before physicists got into the act. New candidates emerged for the elementary constituents of all things. It was established that the thousands of different substances around us were chemically composed of a few elements. The notion also began to get established in the 19th century that each element, when divided up into smaller and smaller pieces would be made of identical "atoms". At that time these atoms were viewed as indivisible entities. The pursuit of elementarity seemed to be reaching its end, with all things in the universe made up of a few species of atoms!

But, as we with out hindsight today could have told them, this state of absolute knowledge of elementarity could not have lasted. With continuing discoveries in chemistry, the number of chemical elements grew larger and larger, each presumably with its own separate atom. The number of such different atoms—several dozens, one per element known was too large for intellectual comfort. The community did not doubt that so many elements, each with its own atom existed. But surely it should not be that the number of basic entities with which to describe the world should be so many in number! So, onwards once again, with further Pursuit of elementarity!

The next level of understanding took a while and involved several major ingredients. On the one hand, a broad pattern had already been noticed in the behaviour of elements, leading eventually to the Periodic Table. If the atom of each element is a separate indivisible entity, distinct from the atoms of other elements, then why should Chlorine behave so similar to Fluorine, or Sodium to Potassium ? The Periodic Table clearly demanded that atoms have an internal structure and that this structure for different atoms must also follow some pattern. With increasing



experimental sophistication in Physics, the inside of atoms began to be probed and it was realised through Rutherford's experiments that atoms themselves were made up of smaller entities, namely a tiny nucleus surrounded by orbiting electrons. The coming of Bohr's model of the atom, Quantum Theory, and the Pauli Exclusion Principle established the shell-like structure of the electronic orbits, and provided the sought-after explanation of the Periodic Table. Soon thereafter it was also realised that the nuclei at the centre of different atoms were themselves made of neutrons and protons. Once again a stage was reached when the pursuit of elementarity seemed to have ended! All things seemed to be made of a small number of elementary constituents, namely  $n$ ,  $p$ ,  $e$ , gamma and possibly the postulated ghostly neutrino.

Again this happy state of affairs proved to be temporary. New subatomic particles continued to be discovered both in cosmic rays and accelerators. By 1960 nearly 30 so called elementary particles were discovered—a veritable zoo and far too large a number to be aesthetically acceptable as the ultimate constituents of all things. Thus once again this cycle—the discovery of a new set of elementary particles at a smaller and more fundamental level, only to find, through more sophisticated experimentation, that this new set also grew in number—was repeated.

Around this time however, some qualitatively new ideas emerged which threatened to alter the very concept of what elementarity means. The rest of my talk will be devoted to explaining these new ideas which have broadened our concept of compositeness beyond naive notions of the "whole" and its "parts".

The time-honoured concept of elementarity which had served well through the ages, all the way upto the middle of this century, was based, as I have already said, on familiar macroscopic day-to-day experience, exemplified by models such as houses and bricks or fruits and their seeds. The house, a composite, is made up of bricks which are its constituents. As behoves a constituent, the brick is smaller in size and in mass as compared to the house. Indeed we all believe that the size and weight of a house will be the sum respectively of the sizes and weights of its constituents, viz., the bricks, the cement, the doors etc. There is no ambiguity about which are the constituents and which is the composite. Obviously, it is the house which is made of the bricks and not vice-versa ! Finally although we have to break open the plaster on the walls to get at the bricks (or break open the fruit to get at its seeds) we know that the

bricks are inside the wall ( and the seeds inside the fruit) even when the latter are not broken open. These are different facets of our intuitive notion of "parts" constituting a composite "whole" and seem too self-evident to warrant a mention. As the pursuit of elementarity by chemists and physicists went deeper and deeper, we learnt that all matter is made of atoms, that atoms in turn are made of nuclei and electrons and that these tiny nuclei are in turn made of protons and neutrons. Throughout this process spanning chemistry, atomic physics and early nuclear physics our intuitive ideas of elementarity, of constituents and composites, continued to hold. Clearly a chunk of iron was much larger and heavier than the individual iron atoms that sit inside it. That iron atom in turn is larger and heavier than its nucleus or the electron and finally the nucleus is larger and heavier than its individual neutrons and protons.

These simple, macroscopic and intuitively self-evident notions of composites and their constituents, however, had to be modified at the next stage in the pursuit of elementarity. This stage began in the late 1940's after the era of atomic physics. By this time the study of the basic constituents of the universe had been given the status of a separate branch of physics called Elementary Particle Physics. Once it was clear that atoms are made of electrons and nuclei and further that the nuclei were made of protons and neutrons, the next step was to look inside these neutrons and protons and see if they were truly elementary or whether they in turn had further internal structure.

To see why, at this sub-nuclear level, the very notion of elementarity may have to be modified, let us consider how sub-nuclear particles (today's elementary particles) are studied. The only way to look at the insides of protons and neutrons is to strike them with a high energy beam of other particles and see what comes out. This is true already at the level of looking at atoms. We can't see atoms by looking at them using ordinary light, even with the most powerful microscope. The reason is that the size of an atom, about 1 Angstrom, is much less than the wavelength of ordinary light, which is a few thousand angstroms. It would be like trying to feel the pebbles on the road while riding on a tractor. A child's tricycle is more likely to feel their effect. In much the same way, to look at atoms we must use light whose wavelength is somewhat smaller than the size of atoms. Such "light" is called an X-ray and it is well known that the atomic structure of solids and various chemical and bio-molecules is done by X-ray crystallography. Equivalently, instead of X-ray photons, one can use a beam of electrons of similar wavelength as is done in electron

microscopy. (Recall that quantum theory associates a wave nature to electrons as well). Either way, whether we use an X-ray beam or an electron beam to look at atoms, the momentum of the X-ray photon or the electrons in this beam will be about ten thousand times larger than that of photons of ordinary visible light, because the wavelength of the former is about ten thousand times smaller. (Recall the famous formula  $E = h\nu = hc/\lambda$ ). Extending the same argument, in order to probe the insides of neutrons and protons which are a hundred thousand times smaller than the atoms, we need waves of wavelength about  $10^{-13}$  or  $10^{-14}$  cm. These are provided by Gamma-rays (mere X-rays will not do) or equivalently by high energy electron or proton beams. These beams are accelerated by giant accelerators across effectively billions of volts. (The giant accelerator at Fermilab in the U.S. delivers a beam accelerated by an effective voltage of nearly a thousand billions volts. The story of these giant accelerators is fascinating in it own right, but we can't go into that here). Not only are the particles in the beam of very high energy, but so are the particles that come flying out when this beam hits the target of a proton or a neutron. Thus the whole process of studying the deeper structure of elementary particles involves dealing with very high energy particles travelling at speeds very close to the speed of light. This is why the name High Energy Physics has become synonymous with particle-physics. By striking protons and neutrons with such high energy beams, a whole host of new particles were discovered in the fragments that flew out. Some of these new particles were also seen in Cosmic rays, which too consist of very high particles. The elementary particles thus discovered included, apart from the Pi-meson, and the Mu-meson, the different Kaons, the Rho, the Omega meson, the Eta, the Lambda, the three Sigmas, and so on. The list became quite long and one was running out of Greek and Roman alphabets in trying to name them ! All this was in addition to the proton, the neutron, the electron, the neutrino and the photon, which were already known. In this way, by 1960 the set of elementary particles had grown into a whole zoo. It was not clear which, if any, of these particles was more elementary than the others and which could be considered as the building blocks of the others.

It was in trying to make sense of this zoo that theoretical physicists of the more thoughtful variety realised that the very notion of elementarity may have to be modified, or even abandoned. There were three ingredients in sub-nuclear physics which combined to force this major qualitative change in our notion of elementarity from the earlier simplistic view.

These three ingredients were the theory of relativity, quantum theory and the presence of the so-called strong interactions—very strong nuclear forces that most of the elementary particles exert on one another. Let us first see why these three ingredients unavoidably enter particle physics.

That Einstein's special theory of Relativity would play a vital role is evident from our earlier remarks that in sub-nuclear physics the participants mostly travel at speeds close to " $c$ ", the velocity of light. The theory of relativity was known since 1905, but its predictions differ significantly from those of the old Newtonian theory only for systems with speeds near  $c$ . The electrons in an atom and the protons and neutrons in the nucleus move at speeds well below  $c$ . By and large therefore the theory of relativity did not play such a crucial role in atomic and nuclear physics.

At most it provided small "relativistic corrections" to the answers. Upon moving into sub-nuclear particles physics however, the ideas of relativity became unavoidable and central to our understanding.

It was already known that neutrons and protons exerted a very strong force on one another, much stronger than the forces involved in atomic physics and chemistry. A majority of the newly discovered elementary particles were also observed to exert similar strong forces on each other. Finally the need to treat elementary particles by quantum principles is evident. Tiny objects can be understood properly only by quantum theory. Even atomic and nuclear physics are couched in quantum theory. Its role in sub-nuclear particle physics, is unquestionable. Thus quantum principles and strong forces were already part of physics at the nuclear level. But when combined with relativity in sub-nuclear particle physics, they formed a very potent mixture which apart from other things, threatens the very notions of elementarity, of the "whole" and its constituent "parts". Let me try to explain how this happened.

Let me begin with an important consequence of quantum theory. In the macroscopic day-to-day world, in order to see what is inside a sealed box, we break it open and see what falls out. Those would be the contents or the "constituents" of the box. Correspondingly, to "see" what is inside an elementary particle we have to strike it with another particle of very high energy, as we earlier argued, and see what comes out. Suppose for example we study the structure of a proton by striking it with a gamma ray photon of 100 Mev energy and look at what comes out. The first complication comes just from quantum theory. As is well known, thanks to the wide publicity that the Uncertainty Principle has received, quantum

theory is not a fully deterministic theory. The outcome of individual interactions cannot be precisely predicted even if you did a complete calculation using all inputs. All that can be predicted are the probabilities of different results. As applied to our experiment of striking a proton with a gamma-ray, various different sets of products may emerge from such a collision. The equations below indicate some of the possible results:

$$(i) \quad \gamma + p \rightarrow n + \pi^+$$

$$(ii) \quad \gamma + p \rightarrow p + \pi^0$$

$$(iii) \quad \gamma + p \rightarrow \Lambda^0 + \kappa^+$$

and so on.

Of course, in a single collision event only one of these outcomes can happen, say,  $n + \pi^+$ . But if the experiment is repeated under identical conditions, the next time  $p + \pi^0$  may emerge from the collision, another time  $\Lambda^0 + \kappa^+$ , and so on. According to quantum theory, the best we can ever do is to predict the probabilities of the different collision-products listed above. This is not just a limitation of the theory. The real world itself seems to carry this probabilistic feature. Repetitions of that collision process with identical gamma rays do yield at different times different products. It is like looking repeatedly at a glass bottle containing some marbles, and finding that every time we look, it contains marbles different in number, size and shape!

From this what are we to conclude ? Is the proton made of a neutron and a  $\pi^+$ , or is it made of a  $\Lambda^0$  and a  $\kappa^+$ ? The answer is that in general we may have to abandon the notion of an object containing a well specified set of parts! The best we can do is to give the probabilities of its having different sets of parts! There is some probability that the proton is made of a  $n$  and a  $\pi^+$ . There is also some probability that it is made of a  $\Lambda^0$  and a  $\kappa^+$ , or a  $p$  and a  $\pi^0$ ! In short as part and parcel of the overall probabilistic nature of quantum theory, the notion of the internal contents of an object also has to be modified into a probabilistic one. (Why does this not happen in the day-to-day world, or even in atomic physics ?) After all quantum theory is supposed to hold universally, and certainly at the atomic level! Yet we say with great confidence that a hydrogen atom is made of one proton and one electron going around it, nothing more, nothing less. The reason is that the probabilities happen in that case to be overwhelmingly in favour of that set of constituents. Strictly speaking there is a tiny but non-zero probability that the hydrogen atom also "contains" an extra  $e^+ - e^-$  pair

in addition to a  $\pi$  and an  $e^-$ . But this probability is so small that we usually neglect in atomic and chemical physics.) Even at the sub-nuclear level, if we are lucky, probabilities in some cases may overwhelmingly favour one specific set of constituents. But it is clear that in principle, we must be prepared to resign ourselves to the notion that if the contents of a particle are to be decided by what comes out of it upon bombardment, such contents may be determined only upto a probabilistic level.

This is bizarre enough, but there is more to come, this time because of relativity. Everyone has heard of that famous equation from relativity theory,  $E = mc^2$ . It is one of the most famous equations in all of science, and is the quantitative statement of the principle of the equivalence and interconvertibility of mass and energy. The symbol  $c$  once again stands for the velocity of light. Consider a particle with some mass  $m$  at rest. When it moves it acquires kinetic energy. Then according to relativity, its mass also increases by an amount equal to this kinetic energy divided by  $c^2$ . Of course light travels extremely fast ( $c = 3 \times 10^{10}$  cm per second) and the increment in mass  $\Delta m = \Delta E / c^2$  would be hardly noticeable if the particle moved at ordinary speeds. But, for the very high energy particles involved in elementary particle physics, the increase in mass due to kinetic energy can be considerable. For instance, a proton coming out of the Fermilab accelerator at 800 gev energy has a mass of about 850 times its normal rest mass. By the same logic, if such a proton slowed down to rest its kinetic energy can be used to create about 400 pairs of other protons and anti protons ! The same is true of a gamma ray photon of similar energy. Furthermore, the principles of quantum theory, explicitly permit such "creation" of particles by using the energy provided by other particles. In short, a high energy particle, whether it be a proton or electron or a gamma-ray photon can slow down as a result of a collision, and in the process create other particles which were not there to begin with. Indeed this is how most of the elementary particles known to us were created in the laboratory and in cosmic rays.

Given this, suppose we study the proton by striking a gamma ray at it, and suppose the result is

$$\gamma + p = n + \pi^+$$

What are we to conclude ? Were the neutron and the  $\pi^+$  meson "inside" the proton before the collision, so that we may take them to be constituents of the proton ? Or were they simply created afresh in the process of collision by using up the energy of the gamma-ray photon ?

There is no independent way of "looking" at the inside of a proton without shining on it light of appropriately small wavelength (i.e. gamma-rays) or some other equally energetic projectile. The possibility that what you see coming out was created in the process of "looking", and not originally present in the target can never be ruled out. Indeed to the best of my knowledge there is no measurable scientific way to distinguish between the two interpretations. In short, the familiar macroscopic way of finding the content of a box by breaking it open and seeing what falls out, does not work at the sub-nuclear level. The very act of "looking" or breaking open can create new particles which might not have been originally "in there". This wipes out an important aspect of our normal notion of "contents".

Next, let us explore the impact of the very strong sub-nuclear forces on the concept of elementarity. When two particles are bound to one another by an attractive force, we know from simple physics that they have a negative potential energy which is larger than whatever kinetic energy they may have. Their "Total energy" in the old pre-relativistic language is negative. According to relativity theory, this corresponds to a reduction in their combined mass as compared to the sum of their individual separate masses. This effect may be represented by the equation

$$\text{Mass (AB)} = \text{Mass(A)} + \text{Mass(B)} - (\text{their binding energy})/c^2.$$

This equation always holds. For instance the combined mass of the earth-moon pair is, strictly speaking, less than the sum of their individual masses because they are bound to each other by the force of gravity. But in that example the reduction in mass is negligible as compared to their gigantic individual masses. Coming down to the atomic scale, the mass of the hydrogen atom is less than the sum of the proton and electron masses by about  $13.6 \text{ eV}/c^2$ . This is still negligible as compared to the proton mass which is about  $940,000,000 \text{ eV}/c^2$ . But when you get to the nuclear level these effects become quite significant. A Uranium nucleus, which is made of 238 nucleons (a common name for neutrons and protons) has a mass roughly equal to what just 236 or 237 separate nucleons would have had ! This is because of the very strong attractive nuclear force (the so-called strong interactions) that binds nucleons to form nuclei. At the sub-nuclear level, similar and possibly stronger forces exist. In principle this can lead to such possible results as A being a bound state of B and C while at the same time being lighter than B! The negative energy of binding, if large enough, can do such havoc. Thus a proton could be viewed in some sense as a composite of a neutron and a  $\pi^+$  meson. What is worse, at the same

time, the same neutron can be viewed as a composite of a proton and a  $\pi^+$  meson!. Such possibilities reduce conventional notions of composites to shambles!

We are not saying that such unfamiliar possibilities necessarily occur in nature. They are however permitted within our modern conception of physics broadened by relativity and quantum theory. In fact such possibilities were very seriously considered in the so-called Bootstrap theory of elementary particles which held sway in the early 1960's. This theory hypothesised that the majority of elementary particles known then (the so-called Hadrons) were all made up of one another! This was a generalisation of the idea described earlier of a proton containing a neutron and at the same time vice-versa. I consider this to be a very brave and revolutionary hypothesis in physics. If subsequently this bootstrap theory was abandoned it was not because anything was conceptually wrong with the idea, but rather because of calculational difficulties and because newer experimental findings led us towards other theories. Among the newer findings was the discovery of quarks and gluons as the constituents of all hadrons. But these again are not constituents in the normal sense. Part of their properties is that they cannot be removed from their composites and do not occur in isolation. During the past 15 years theorists have gone further. The recent rage is the famous superstring "Theory of Everything". This is a remarkably beautiful and sophisticated theory which is still incomplete. Unfortunately time does not permit us to give a simple introduction to these very recent developments. Suffice it to say that the story of elementarity is far from finished, and it remains to be seen what newer heights of conceptualisation it will bring.

I conclude my lecture by paying my homage to Dr G P Chatterjee, in whose memory this lectureship is endowed.





**Mithan Lal Roonwal** (b. 18 September 1908; d. 22 July 1990) did Ph.D. (1935), Sc.D. (1967) from University of Cambridge, UK. He was Emeritus Scientist (CSIR), Desert Research Station, Jodhpur (1969-73); Professor & Head, Department of Zoology (1965-66) and Vice-Chancellor, University of Jodhpur.

Roonwal did extensive research on insect embryology, ecology, morphology, and taxonomy of locusts and termites; forest entomology; and animal behaviour especially of primates. He put forward: a new theory of gastrulation in insects; new phase characters, polymorphism and gregarization theory in locusts; and new taxa in termites. He made significant observations on primates, especially on intraspecific variation and tail-form and carriage.

Roonwal was Life Fellow of Zoological Society of India; Honorary Member, All-Union Entomological Society, Soviet Academy of Sciences (Moscow), and Entomological Society of India, Member, INSA Council (1961-63). He was the recipient of Tata Medal (Zoological Society of India) (1956); Har Swarup Memorial Lecture Award (INSA) (1984).

*Mithan Lal Roonwal was elected to the fellowship of the Academy in 1945*

# RECENT RESEARCHES ON WING MICROSCULPTURING IN TERMITES (ISOPTERA), AND ITS EVOLUTIONARY AND BIOLOGICAL SIGNIFICANCE

M L ROONWAL

*A brief account is given of recent researches carried out by the author and his co-workers on wing microsculpturing in termites (Isoptera) and its evolutionary and biological significance. Some 80 genera and over 250 species, belonging to all the major families and subfamilies of termites from all parts of the world, have been studied.*

*Hairs are present in moderate numbers, especially on wing margins. Sometimes they are abundant on the membrane as well, at others they may be almost absent. Apart from hairs, at least eight major types of tiny, microsculpturing elements are present in great abundance on both the upper and lower surfaces of wings. According to their structure, etc., these elements have been named as papillae, arrowheads, tubercles, spearheads, pimples, micrasters, microsetae and rods. They vary in size from about 0.5 to 24  $\mu\text{m}$ ; and in density from about 900 to 13000 per square millimetre surface area. On a single wing as many as over half a million of these bodies may be present on one surface, and thus over a million on both surfaces together. The earlier Light Microscope (LM) studies have recently been supplemented with Scanning Electron Microscope (SEM) examinations which have thrown fresh light on the intimate structure of the microsculpturing elements.*

*The evolutionary and biological consequences of the presence of these elements are discussed. The enormous dead weight which they add to the wings, and the great rugosity that they create on the wing surfaces must greatly impede the aerodynamic efficiency of wings and slow down flight (termites are known to be slow and erratic fliers). This, in turn, prevents dispersal to distant and perhaps hostile places by intrinsic powers alone. The dispersal of some widespread species has occurred by extrinsic means such as driftwood, commerce, heavy windstorms, etc. The slow flight of swarms and their early descent to the ground make them easy prey to predators on the wing such as birds and bats and on the ground to various reptiles and mammals, including man. Microsculpturing has also proved to be of considerable utility in the taxonomic differentiation of termites especially at the generic and*

*specific levels Comparisons are made with the allied insect orders Dictyoptera (cockroaches, etc.), Zoraptera and Embioptera.*

## INTRODUCTION

### 1. General

I am greatly honoured in being the first recipient of the Academy's Har Swarup Memorial Lecture Award. I had the privilege of having Professor Har Swarup's friendship. He was a fine zoologist who devoted his working life first to the study of fishes, and, in later years, as the Vice-Chancellor of the Vikram University at Ujjain, to educational administration. He was soft-spoken and had a genial temperament and his passing away at a relatively young age is a serious loss to Indian zoology.

In today's lecture I would like to speak briefly about some aspects of my own researches during the last 15 years or so on wing microsculpturing in termites (Isoptera) and its evolutionary and biological significance.

Most insects carry a pair of wings on either side of the body. These wings are thin, double-walled, cuticular structures and they bear a system of thickened streaks called veins which give them strength and prevent them from being torn apart in high wind. Insect wings are generally naked except for a few hairs. But in certain groups they are clothed with a dense covering of hairs (setae or macrotrichia of authors) as in caddis-flies (Trichoptera) and sand-flies (Diptera: Psychodidae). In some primitive groups of holometabolous insects (Zoraptera) both long hairs (setae or macrotrichia) and minute hairs (microtrichia) occur in abundance. In moths and butterflies (Lepidoptera) wings are clothed with a dense covering of tiny, overlapping scales.

### 2. Origin of Termites

Before dealing with termite wings, I should tell you briefly about the origin of termites and the part which flight plays in their lives.

Insects (which today dominate the world and constitute fully two-thirds of the one million and odd known species of animals) first appear in earth's history some 350 million years ago as winged fossils in the Upper Carboniferous rocks of North America. The first termite fossils (wings of *Cretatermes carpenteri* Emerson, subfamily Cretatermitinae, family Hodotermitidae, vide Emerson 1967) have been found in the early Mesozoic Age (Mid-Cretaceous of Labrador, North Canada, about 130

million years ago). The early termites had already evolved the social way of life (as evidenced by the occurrence of polymorphic forms for different functions). This was, in fact, the first social system evolved by animals—millions of years before Man came upon the scene, some two million years ago.

Today, there are about 2000 known living species of termites in the world. Of these, about 270 occur in the Indian Region.

### 3. *Termite Flight*

One of the secrets of the phenomenal success of insects in the modern world is their ability to fly long distances. Familiar examples are locusts which can fly hundreds of miles at a stretch in enormous swarms composed of millions of individuals. Flight of termites are not so spectacular, but they nevertheless play a vital, and in fact an essential, role in termite life-history.

Termite flight occurs periodically, usually once or twice a year, in swarms, depending upon the production of winged imagoes or alates in the colony. In the Indian Desert, for example, swarming occurs once a year and is clearly coordinated with the rainy season, the swarming duration being one to a few weeks. The actual swarming is triggered by the immediately preceding heavy rainfall (within a day or two of swarming), and alates, otherwise fully equipped and ready for flying, may remain idle in the nest for a month or more, awaiting the arrival of a heavy shower of rain (Roonwal 1976). While heavy rain triggers swarming, most termites do not swarm out during heavy rain. If swarming has already begun, it stops during heavy rain, to begin again when the rain stops. Furthermore, each species has its own characteristic hour of swarming—early morning, forenoon, early afternoon, late afternoon, evening and night. (For more details, vide Roonwal 1976, 1979, 1983c.)

A swarming flight is of short duration, hardly an hour or two. The alates soon descend to the ground, and by turning and twisting the wings they cast them off, leaving only stumps. Then usually follows a fast process of courtship during which a male closely follows a female (touching her hind end with his mouth-parts) as she careers around irregularly on the ground. After a while the pair finds a suitable hole in the ground where it settles down for life, to start a new colony. A week or two later, the first batch of eggs, hardly 10 or 20, is laid. In two or three weeks they hatch out into young, whitish larvae which undergo a few moults and

become "adults". The early batches of eggs produce only workers, then a few soldiers are also produced, but the winged or reproductive forms are produced at long intervals, usually only once or twice a year. The workers gradually take on the entire manual work of the colony, such as feeding the parents and the young larvae, building the nest and mound, cleaning the colony, transferring the eggs to nurseries, etc. The breeding pair is often imprisoned for life in a royal chamber made of earth or wood-carton, and is then known as the royal pair (king and queen) whose sole function is to reproduce. A queen may lay an egg per second and, in a life of several years, it may lay millions of eggs. A large termite colony may consist of a few hundred to as many as 3 to 4 million individuals, the population comprising of a royal pair (sometimes a few pairs), up to about 10% soldiers (a few genera are soldierless), and the remainder are all workers.

The reproductive caste in termites possesses, in the adult stage or imago, two pairs of long, thin transparent wings (figure 1) which are subsequential in size and nearly similar in venation—hence the name of the termite order, Isoptera (Greek: *iso*, equal; *ptera*, wings).

In termite literature there is only casual mention of the occurrence of minute structures on wings. They were variously named as nodules, granules, dots, punctations, micrasters, etc. No systematic study was ever made, and there were no descriptions or illustrations. The confusing and often contradictory and vague terminology did not help in understanding their true nature. No mention of microsculpturing is made in the standard accounts of termite morphology by Grassé (1949), Weesner (1969) and Weidner (1970), and of insect morphology in general (Inniss 1965, Chapman 1982).

A detailed and systematic study of these structures was started by me some 20 years ago, and the first account, describing numerous thin, cuticular rods in some Macrotermitinae (*Odontotermes*, etc.) appeared in 1967 (Roonwal and Chhotani). The words which followed were published in a series of papers during the last ten years or so (1974-1985) (vide the references). In them were examined in detail wing microsculpturings in all the major families and subfamilies of termites, represented by some 80 genera and over 250 species obtained from all parts of the world. The nature of these structures and their evolutionary and biological significance are now fairly clear.

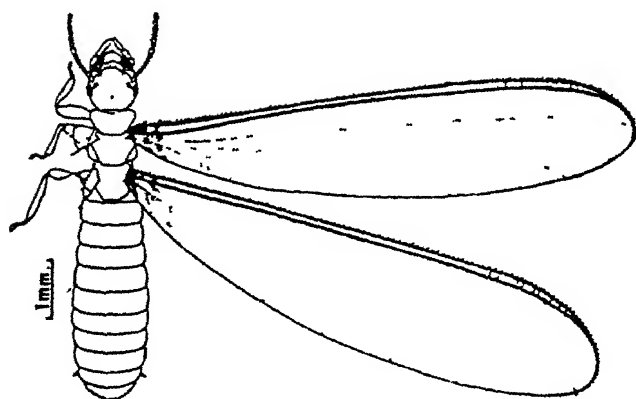


FIG 1 A winged termite (*Eremoterme paradoxalis*), in dorsal view (wings of the left side not shown) Note the large, subequal and similar wings

Recently, I have carried out scanning electron microscope studies which have revealed the intimate structure of these elements.

## RESULTS

### (A) Light Microscopic (LM) Studies

I will first enumerate briefly the result of Light Microscopic Studies (LM) on the morphology of the microsculpturing elements.

- (i) The microstructures occur on both the upper and lower surfaces of wings. Some types are localised in distribution (e.g., the papillae which are, as a rule, confined to the wing margins); others occur all over the wing surfaces.
- (ii) There are at least eight main types of them (apart from hairs) which have been termed as follows, depending upon their shape and general structure (figure 2):
  1. Papillae (figure 2 A-D).
  2. Arrowheads (figure 2 E).
  3. Tubercles (figure 2 F, G).
  4. Spearheads\* (figure 2 H).

---

Spearheads, which occur in some nasutitermitine genera (*Afrosubulitermes* and *Mumeutermes*) and in some species of *Coptotermes* had earlier (Roonwal et al. 1981, Roonwal 1983 b) been regarded as a type of one-armed micrafter. On

5. Pimpules (figure 2 I; plate 11).
6. Micrasters (figure 2 J-L; plates 1-9).
7. Microsetae (figure 2 M)
8. Rods (figure 2 N-P; plate 10).

Further, there are subtypes; for example, micrasters are of at least 10 principal types and several subtypes. The broad distribution of the various types in the families and subfamilies is shown in figure 3.

- (iii) They differ from one another in shape, size, density and location on wings.
- (iv) The various shapes which occur are, as the names imply, as follows: papillae (tongue-shaped or thorny); arrowheads (shaped like > or {}); tubercles (granular humps which are either crescentic or angular); spearheads (longish, subconical and sharply pointed, with the base jagged); pimpules (minute, glassy, rounded or slightly pointed, refractile dots; also flowery shapes, as seen in scanning electron microscopic images); micrasters (1- to 8-armed, with shapes varying as I, V, Y, +, X, K, T, H, etc., the many-armed being often asteroid (\*) in shape; in scanning electron microscopic images, a leafy or flowery form, with a basal vase and 0 to 4 central rods or arms, is also seen; microsetae (thin and hair-like but very small); and rods (thin, cuticular, rod-like structures, sometimes lying subhorizontally, and others subvertically).
- (v) The size varies from about 0.5 $\mu$ m to 24 $\mu$ m (i.e., about 400th to 2000th part of a millimetre).
- (vi) Their density of distribution, per square millimetre surface area of a wing, varies with the type of microstructure as well as the species, and ranges from about 900 to 13,000. Wings of some species carry as many as over half a million of these structures on one surface or over a million on the entire wing (table 1).
- (vii) All the eight types are not present simultaneously in a species. Papillae are present *universally* in all species. Of the remaining seven types, not more than two or three may be present, or none at all, depending upon the species. Some examples are illustrated in figures 4-12. On the basal scale and near it, the sculpturings get distorted (figure 13).

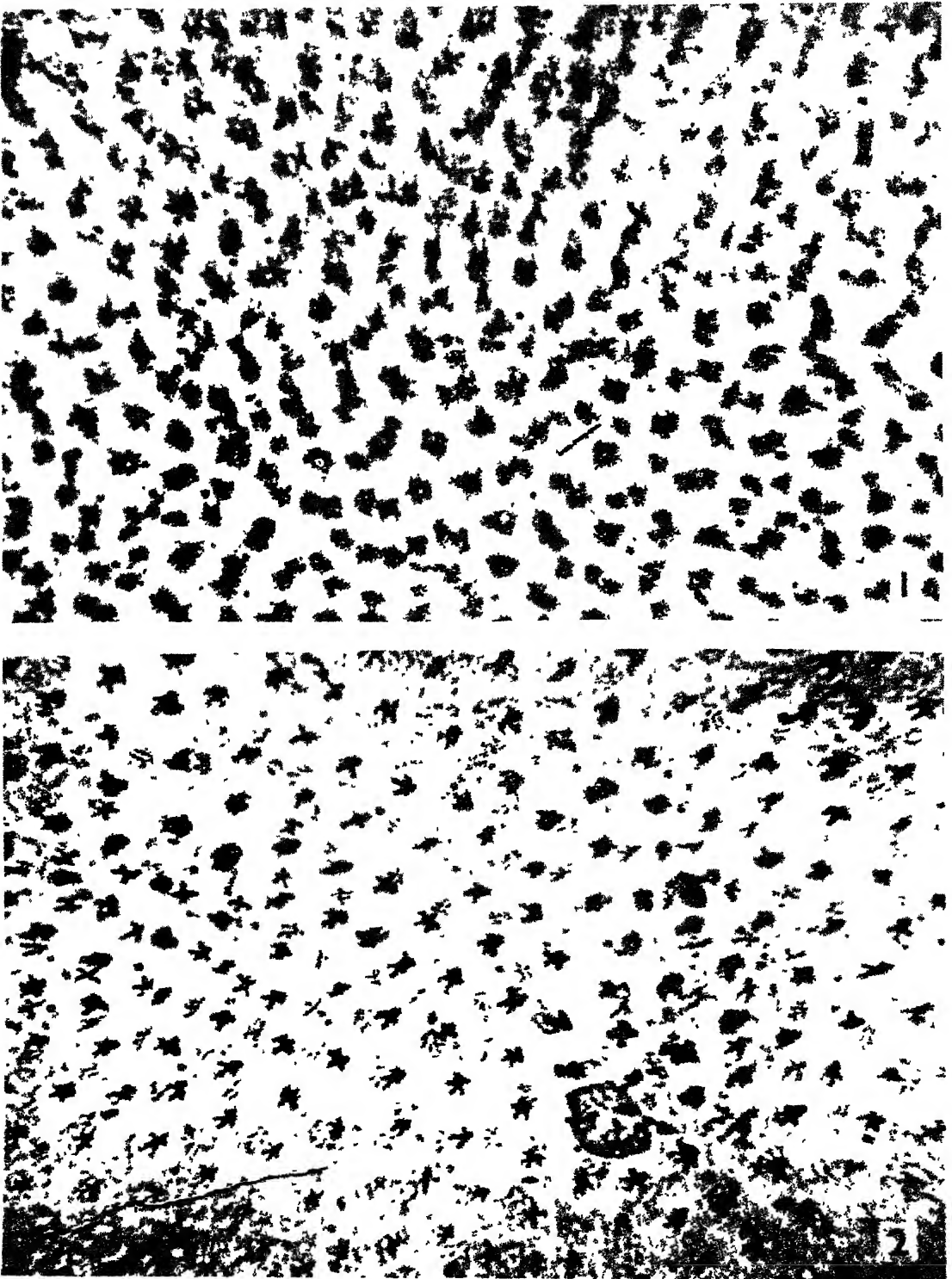


PLATE 1. 1-2, Light Microscope photomicrographs of surfaces of forewings, to show microsculpturing. 1. *Heterotermes indicola* (Rhinoitermidae, Heterotermitinae). Note curved, subcircular arrangement of micrasters which are thick and complex ( $\times 653$ ); 2, *Trinervitermes biformis* (Termitidae, Nasutitermitinae). Micrasters thinner, smaller and simpler



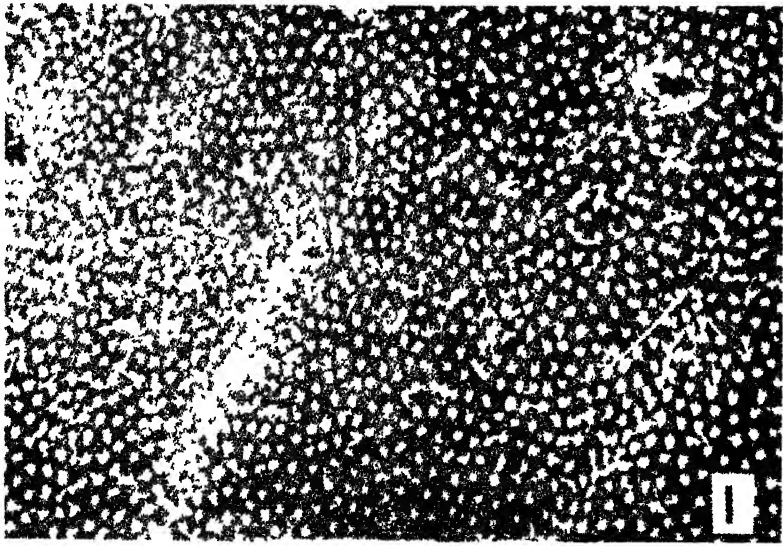


PLATE 2 1-2, *Heterotermes indicola* (Rhinotermitidae, Heterotermitinae). Scanning Electron Micrographs of dorsal surface of forewing, to show micrasters (cf plate 1, figure 1) 1. (x240). 2 More magnified (x2400). Note flowery, vase-like micrasters with pistil-like rods protruding out. Also see plates 3 and 4

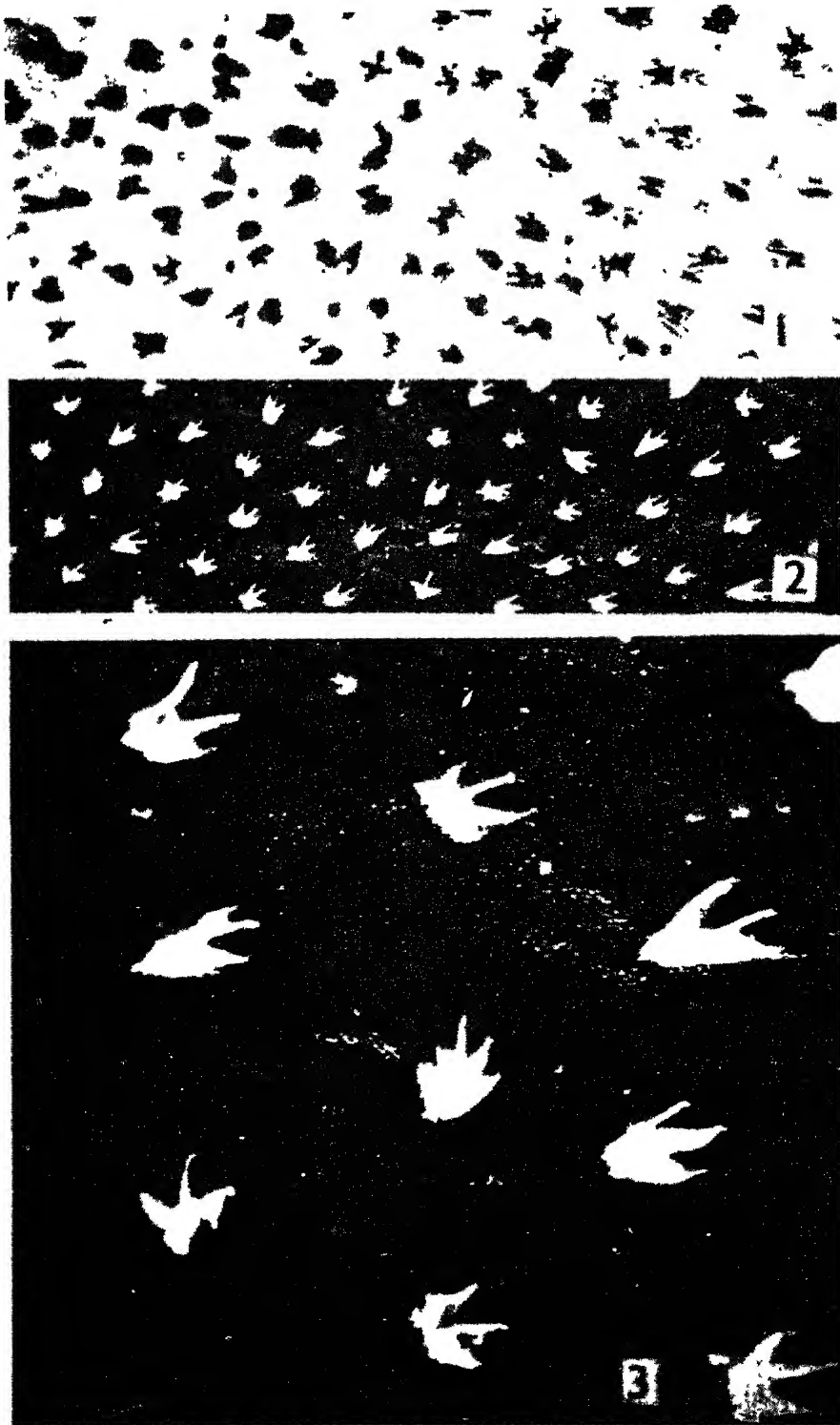


PLATE 3 1-3, *Eremotermes paradoxalis* (Termitidae, Amitermitinae) Views of dorsal surface of hindwing, to show micrasters. (also see Plates 2 and 4) 1, LM photomicrograph. (x700). 2, SE micrograph (x600) 3, Same, more magnified (2400x). Note the flowery, vase-like micrasters which are thinner and simpler than in *Heterotermes* (cf. Plates 1 and 2)

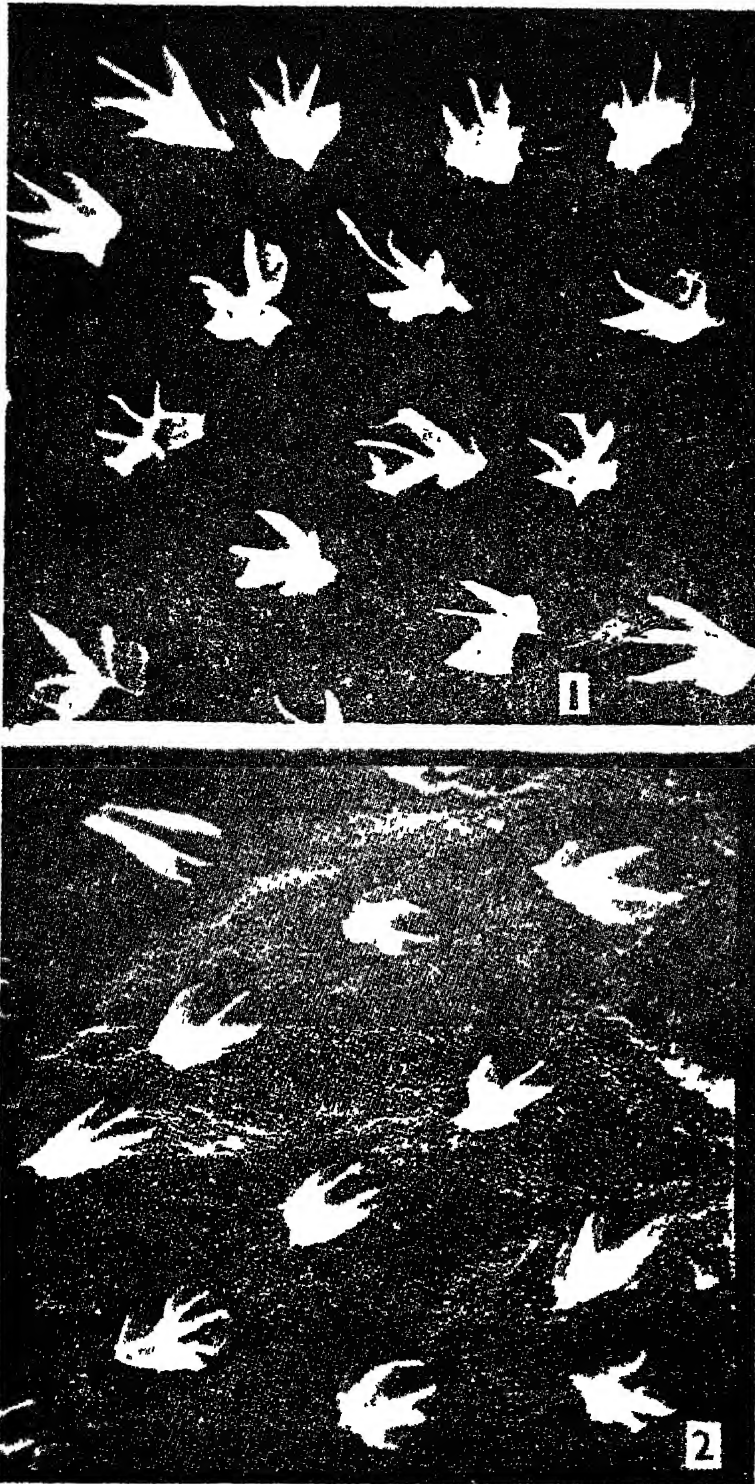


PLATE 4 1-2, SE micrographs of surface of forewings, to show micrasters. (x2400) 1. *Amitermes belli* (Termitidae, Amitermitinae) Micrasters complex. 2. *Eremitermes paradoxus* (Amitermitinae) Micrasters thinner, narrower and simpler

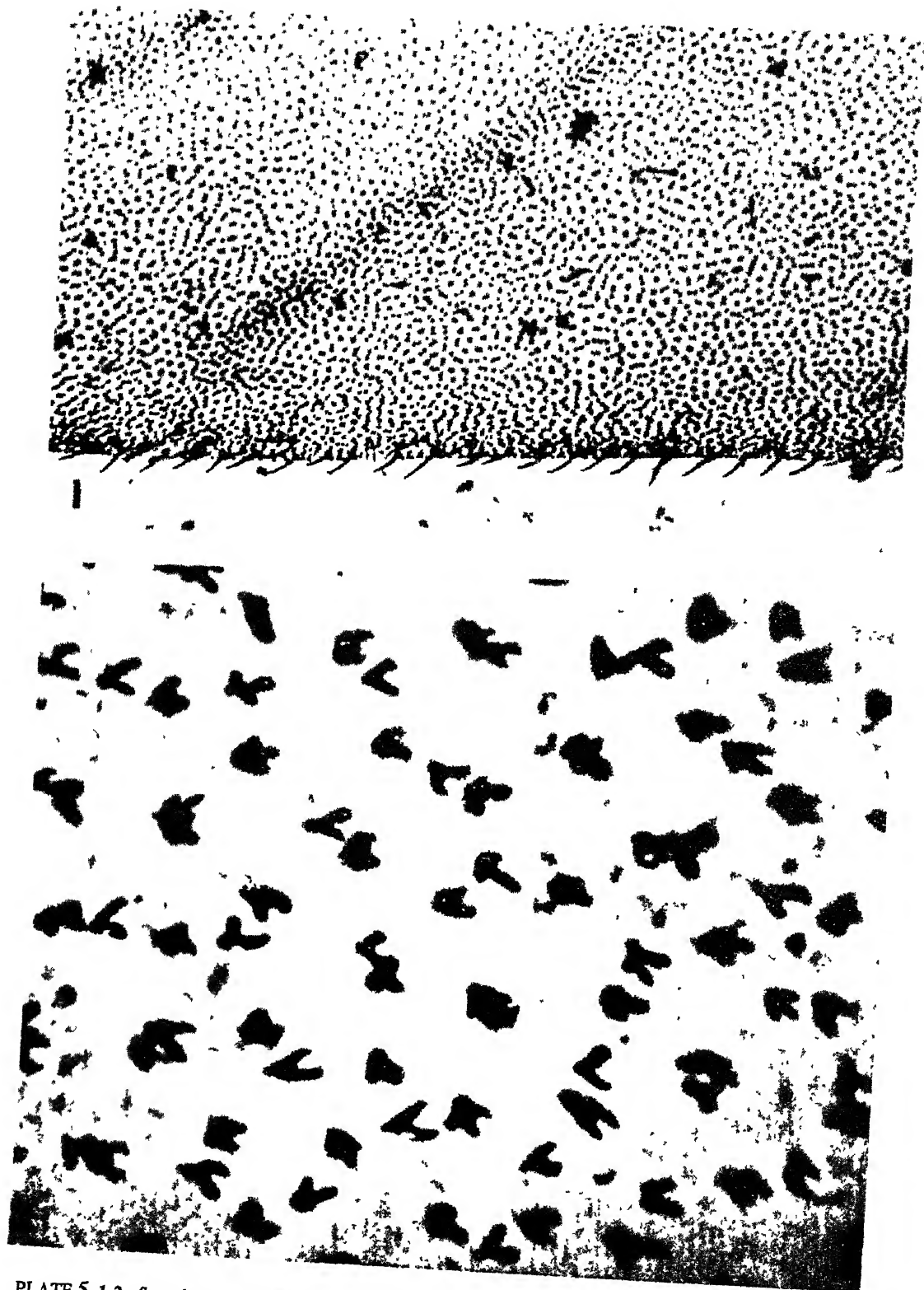


PLATE 5 1-2, *Speculitermes silvestrii* (Termitidae, Amitermitinae) LM photomicrographs of surface of hindwing, to show the simple micrasters. 1, Near posterior margin ( $\times 291$ ). 2, Same, more magnified ( $\times 1117$ )



PLATE 6 1-2, LM photomicrographs of surface of hindwings of two termites (both Termitidae, Amitermitinae), to show the simple micrasters. 1, *Speculitermes sinhalensis* (x1150), 2, *Anoplotermes brevipilus* (x1150)

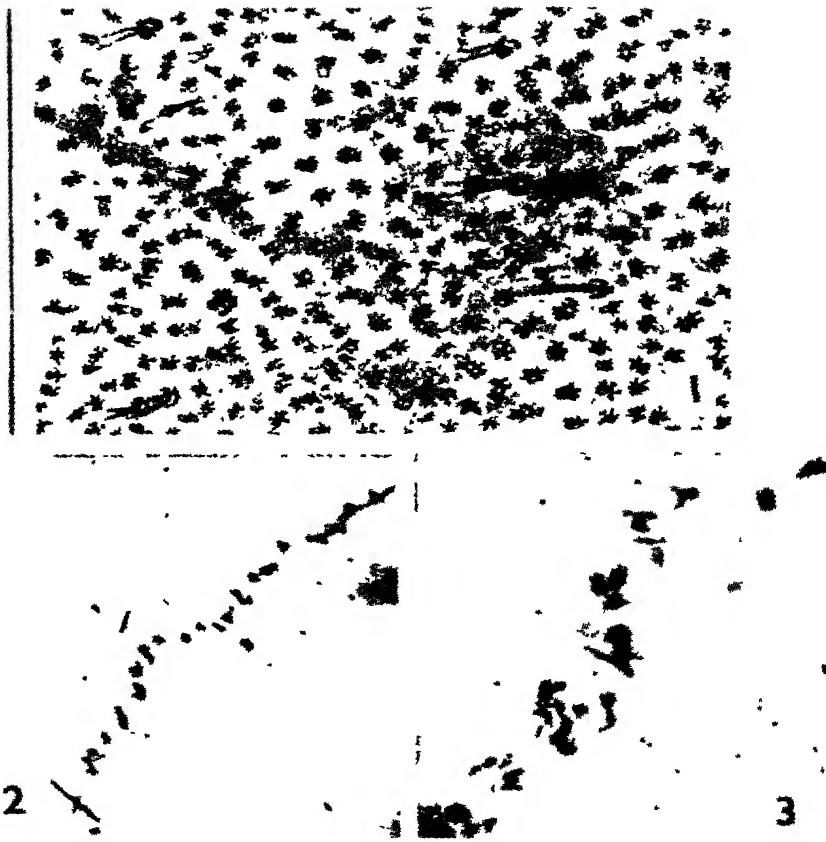


PLATE 7 1-3, *Angulitermes jodhpurensis* (Termitidae, Termitinae) LM photomicrographs of hindwing, to show micrasters and a few hairs 1 Surface view (x700) 2 Transverse section of wing membrane. Note the micrasters (black spots on either side of membrane (x400) 3 Same, more magnified (x1075)

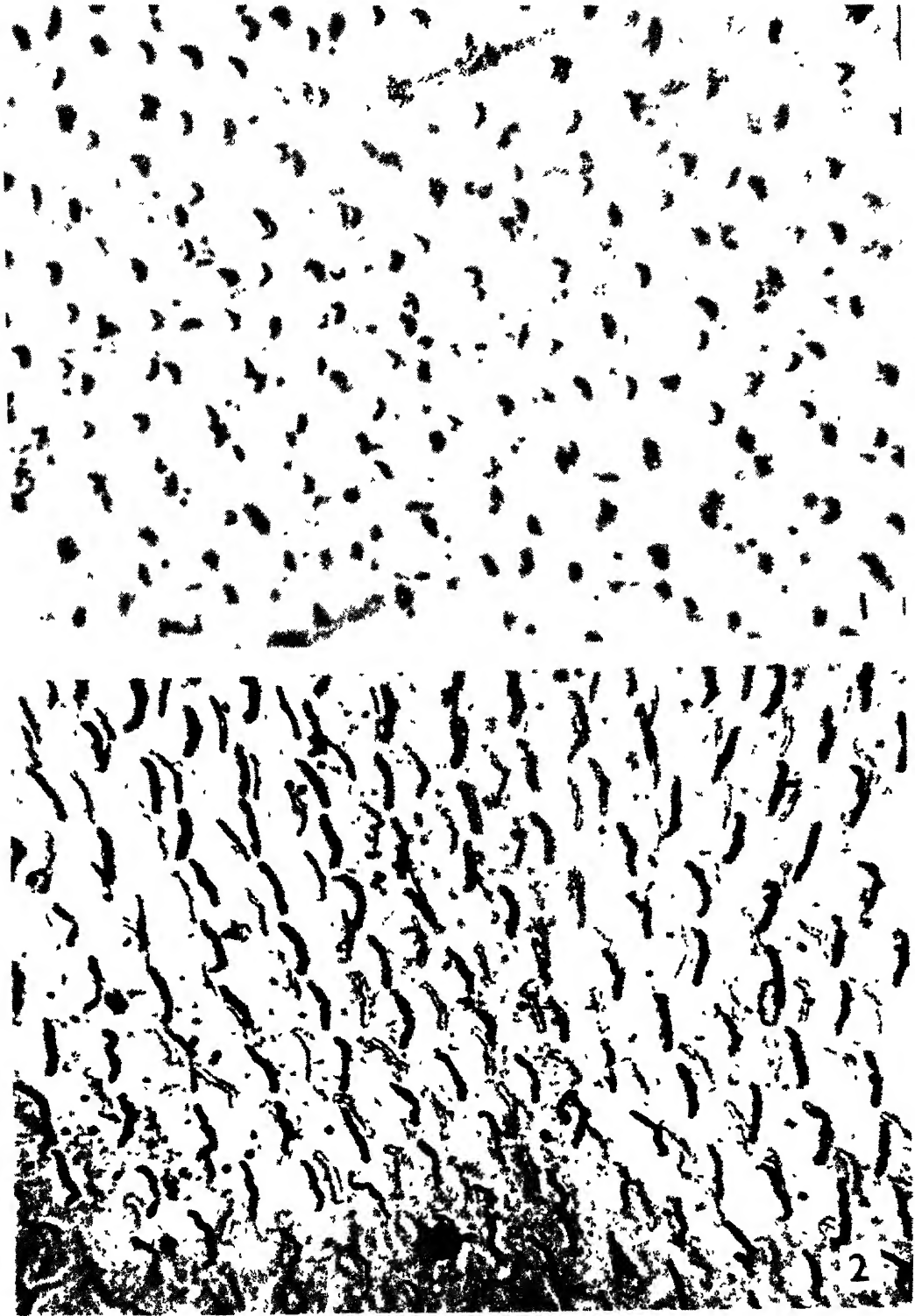


PLATE 8 1-2, LM photomicrographs of dorsal surface of forewings, to show microsculpturing. (x700) *Microtermes obesi* (Termitidae, Macrotermitinae). 1, Small, rod-like bodies which are really simple, vase-like micrastars (cf. plate 9). 2, *Odontotermes obesus* (Macrotermitinae). Longish, subvertical, rod-like bodies (also see Plate 10) (x600).



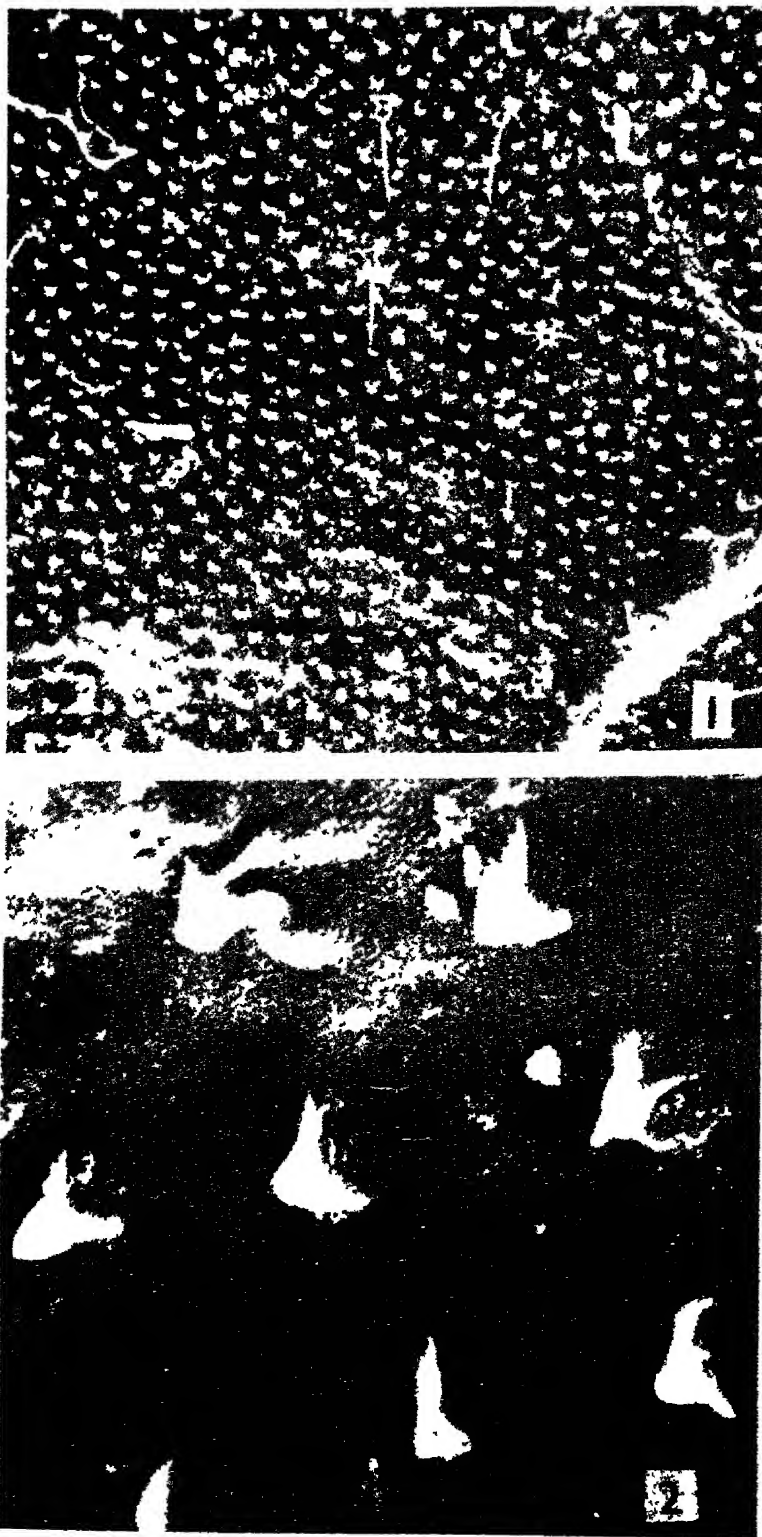


PLATE 9 1-2, *Microtermes obesi* (Termitidae, Macrotermutinae) SE micrographs of dorsal surface of forewings, to show simple vase-type micrasters (which look rod-like under the light microscope (cf Plate 8, figure 1). 1, (x240). 2, Same, more magnified (x2400)



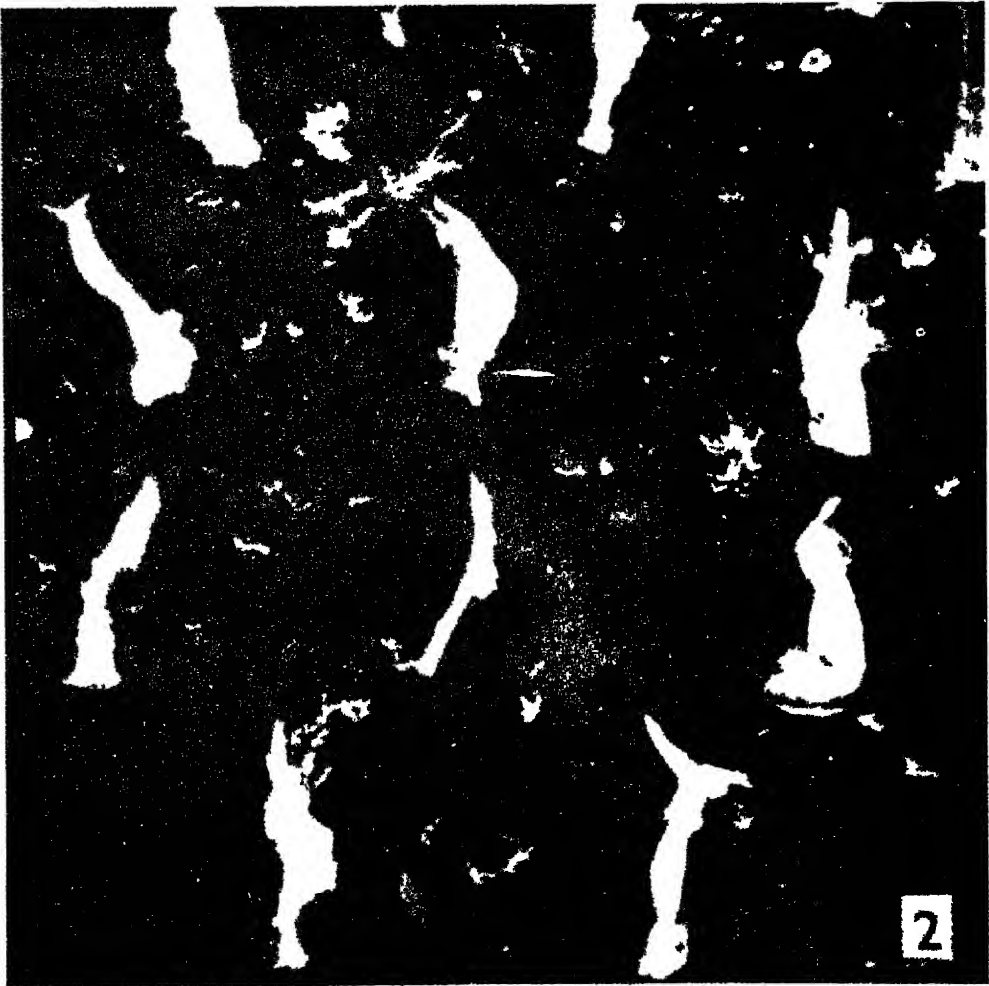


PLATE 10 1-2, *Odontotermes obesus* (Termitidae, Macrotermutinae). SE, micrographs of surface of forewing, to show subvertically disposed rods (cf Plate 8, figure 2) 1, (x600) 2, Same, more magnified (x2400)

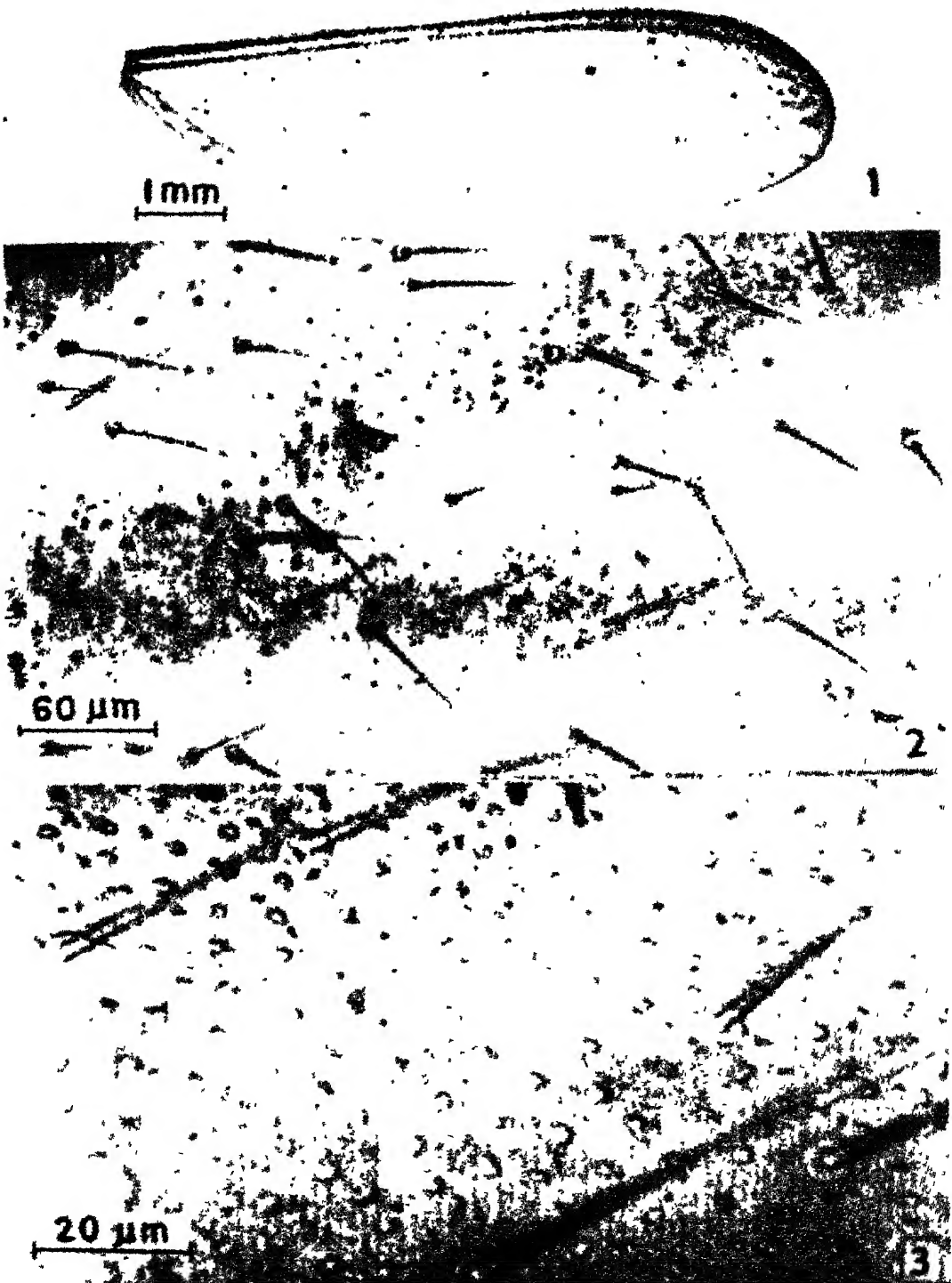


PLATE 11 1-3, *Coptotermes amanu* (Rhinoitermitidae, Coptotermitinae) LM photomicrographs. 1, Right hindwing, without basal scale. The two anterior veins are dark and thick, the other veins are thin and faint. Wing surface covered all over with fine hairs. ca  $\times 13$ . 2, Same, middle of proximal third of wing surface, more magnified, to show hairs and pimples (minute, dot-like bodies) all over ( $330\times$ ). 3, Same, near wing suture, to show hairs and pimples. ( $\times 1200$ ). (Ex Roonwal et al. 1979b)



PLATE 12 1-2, LM photomicrographs of surface of forewings 1, *Coptotermes heuni* (Rhinoitermitidae, Coptotermitinae), to show hairs and pimples ( $\times 700$ ). 2, *Syntermes territus* (Termitidae, Nasutitermitinae) Wing densely covered with hairs (density ca.  $1000/\text{mm}^2$ ) ( $\times 305$ )

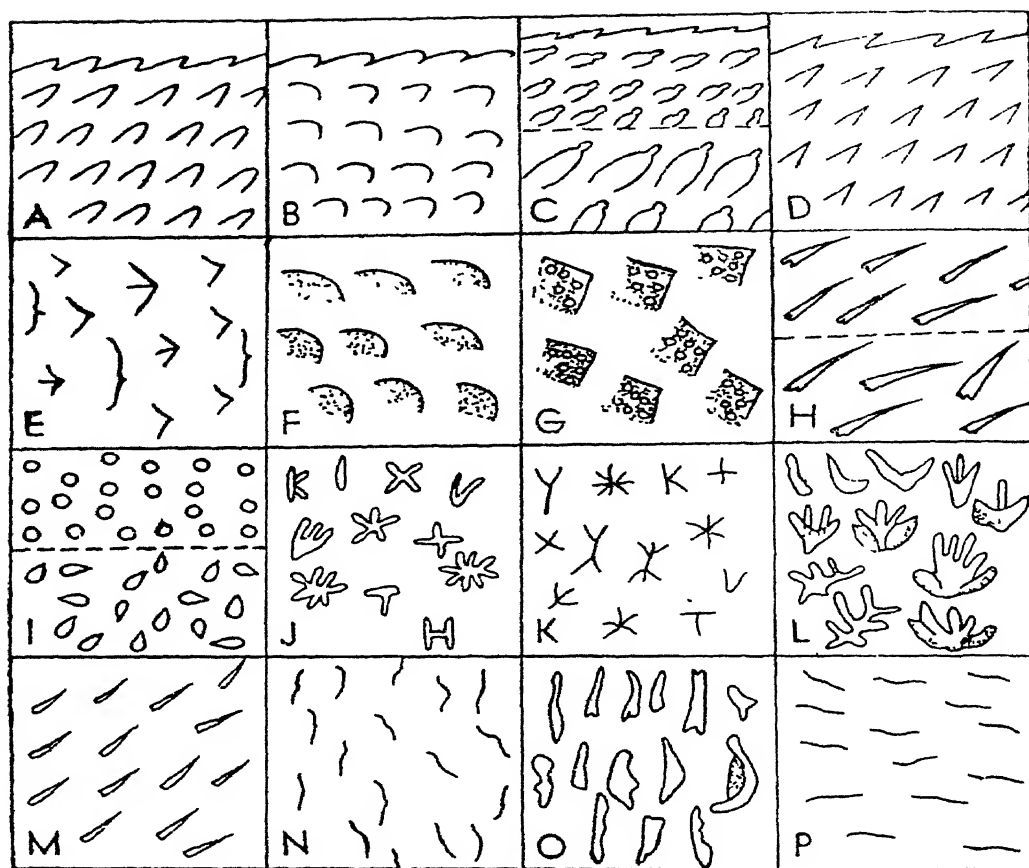


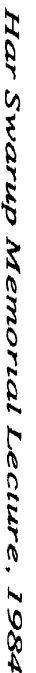
FIG 2 A-P Principal types of microsculptures on wing surfaces in termites (light microscopic images, except where otherwise stated) A, Papillae, finger-shaped or tongue-shaped; B, papillae, subcrescentic, C, Papillae, nipple-shaped (lower half greatly magnified), D, Papillae, pointed, thorny or spiky, E, Arrowheads, F, Tubercles, with crescentic margin; G, Tubercles, with angular margin; H, Spearheads Upper half, smaller and with incurved base; lower half, larger and with the incurved base having a swelling in middle; I, Pimpules (upper half rounded, lower half subconical); J, Micrasters, thick type; K, Micrasters, thin type, L, Micrasters, as seen in scanning electron microscope images; M, Microsetae, N, Rods, subvertical; O, Same, scanning electron microscopic images; P, Rods, subhorizontal. (Note: Papillae occur universally in all species. Of the remaining types, not more than two or three are present in a species, or none at all )

- (viii) Papillae are generally localised in distribution (being confined, in one to several rows, on and near the wing margins), while other structures are more widely distributed. The overall impression of the pattern is that of a heavily embroidered *sari* of an Indian lady, and the similarity includes even the borders (formed by rows of papillae on the wing margins).
- (ix) Microstructures can be grouped into two broad categories, viz., directed and non-directed. The directed ones are the papillae, arrowheads, tubercles and spearheads which point towards the distal end of the wing. The remainder do not display any marked orientation.
- (x) Micrasters, which show the greatest range of variation in structure, also show considerable differences among species of the same genus. Thus, in *Microcerotermes* some species may possess only 3 types (III to V), while others have the whole gamut of 10 types (I to X).

**Table I**  
*Number of microsculpturing elements on wing surfaces of some termites*

| Species   | Total area of one surface of wing (excluding scale) | Micro-sculpturing elements | Density of micro-sculptures (per mm <sup>2</sup> ) | Total number of micro-sculpturing elements |                               |
|---|---|----------------------------|--|--|-------------------------------|
|   |   |                            |  | On one surface (dorsal)                    | On both surfaces (calculated) |
| 1 <i>Heterotermes gertrudae</i> (Rhinotermitidae) | 16 mm <sup>2</sup> (forewing)                       | Micrasters                 | 6000   | 96000                                      | 192000                        |
| 2 <i>Speculitermes sinhalensis</i> (Termitidae)   | 27 mm <sup>2</sup> (hindwing)                       | "                          | 3850   | 103950                                     | 207900                        |
| 3 <i>Eremotermes paradoxalis</i> (Termitidae)     | 12 mm <sup>2</sup> (forewing)                       | "                          | 5670   | 68040                                      | 136080                        |
| 4 <i>Angulitermes jodhpurensis</i> (Termitidae)   | 13 mm <sup>2</sup> (hindwing)                       | "                          | 8100   | 105300                                     | 210600                        |
| 5. <i>Odontotermes obesus</i> (Termitidae)        | 137 mm <sup>2</sup> (forewing)                      | Rods                       | 4520   | 619240                                     | 1238480                       |

## 1964



*A*, arrowheads, *m*, simple (nonasteroid) micrasters, *M*, complex (asteroid) micrasters, *Pm*, pimples, *Pp*, papillae, *R*, rods, *S*, microsetae, *T*, tubercles

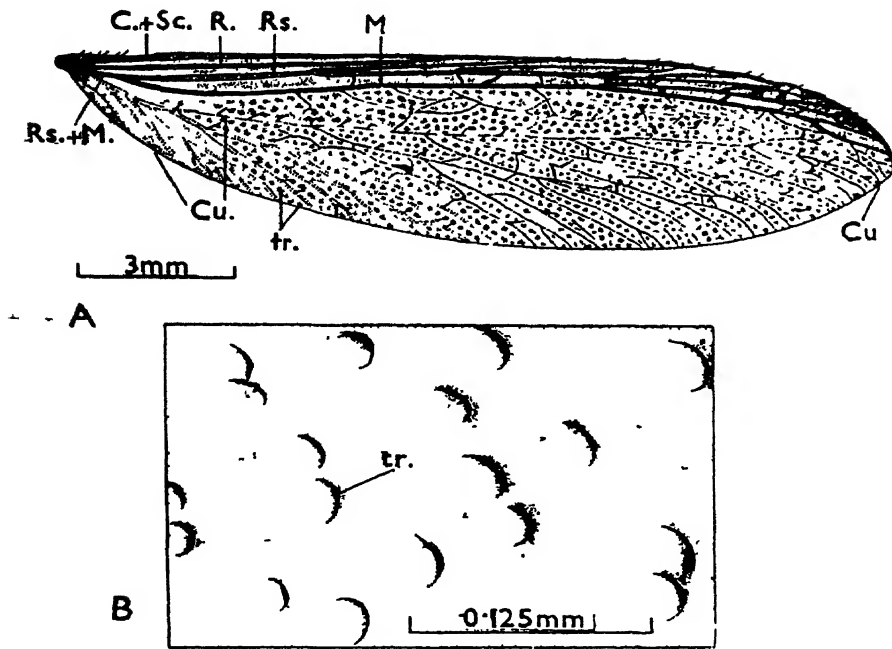


FIG 4 A-B *Postelectrotermes militaris* (Kalotermitidae) A, Hindwing, showing venation and microsculpturing (tubercles), B, Same, central part of wing surface more magnified  
 C + Sc, costa+subcosta, Cu, cubitus, M, media, R, radius, Rs, radial sector; Rs+M, common stem of radial sector and media tr, tubercles

#### (B) Scanning Electron Microscopic (SEM) Studies

(figures 14-15; and plates 2-4 and 9-10)

Recent Scanning Electron Microscopic studies (Roonwal 1985, and in press) have thrown much fresh light on the intimate structure of the microsculpturing elements. For example, structures called micrasters, which in LM images appeared to be plain Vs with simple arms, have, under SEM, proved to be sheath-like or flower-like bodies with a basal vase with either 1 to 4 central arms (or none) protruding out. Wavy, sinuous rods (*Odontotermes*, etc.) do not appear to be really sinuous, and the apparent sinuosity under LM is really due to the varying irregular thickness of the rods. The arrowheads which appear to be smooth under LM, show a series of large teeth or serrations on the outer margins.

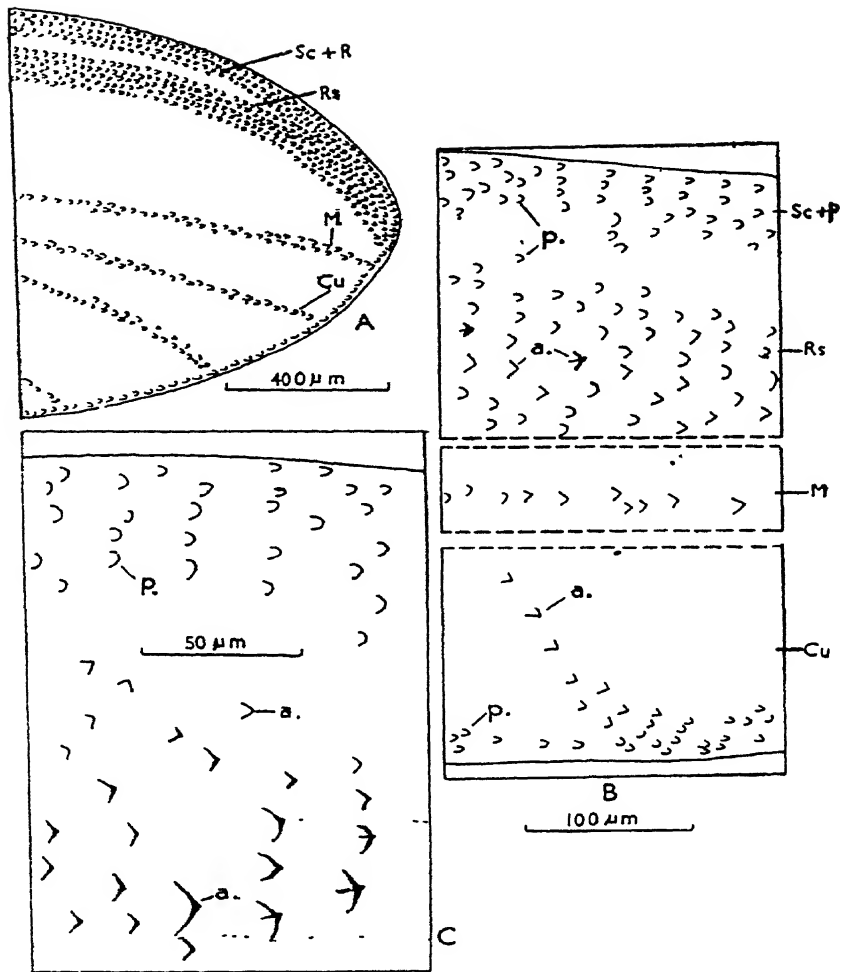


FIG 5 A-C *Stylotermes dunensis* (Stylotermitidae) Portions of distal parts of right hindwing, to show microsculpturing (papillae and arrowheads), A, In low magnification; B, More magnified, to show arrangement of papillae and arrowheads in different parts of wing (upper, anterior margin, middle, across media, lower, posterior margin; C, Anterior edge, more highly magnified  
 a, arrowheads (small and large), Cu, cubitus, M, media, p, papillae; Rs, radial sector, Sc + R, subcosta + radius



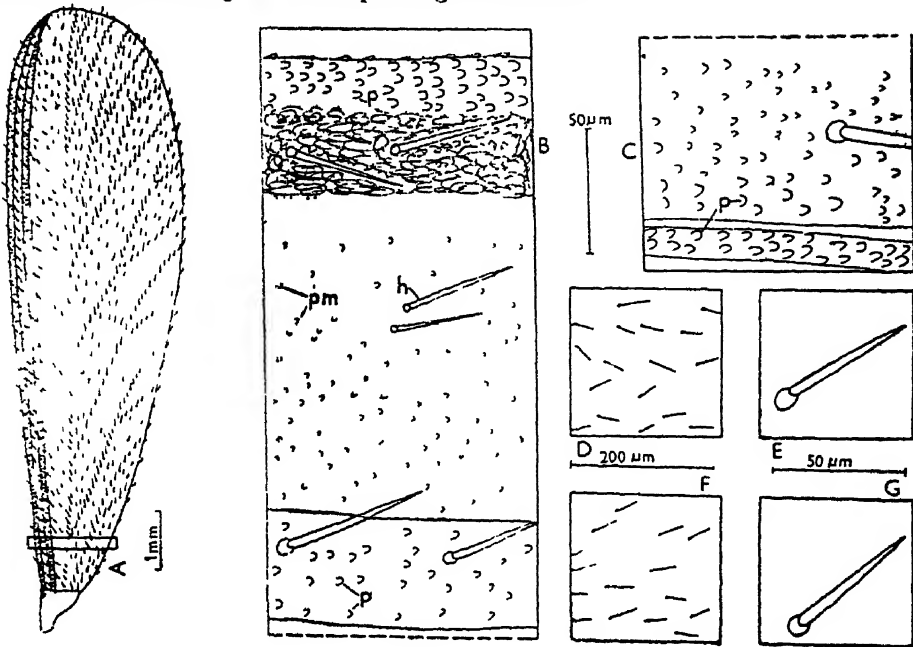


FIG 6 A-G *Coptotermes acinaciformis* (Rhinotermitidae, Coptotermitinae) Wings and their parts, to show hairs and microsculpturing (A-E, forewing, F, G, hindwing) A, Right forewing, note the hairs all over The rectangle indicates the proximal portions enlarged below, B, Anterior margin, C, Posterior margin, D, Middle part (forewing), E, Same, a single hair more magnified, F, Middle part (hindwing), G, Same, a single hair more magnified (Roonwal et al 1979b) k, hairs, p, papillae, pm, pupules

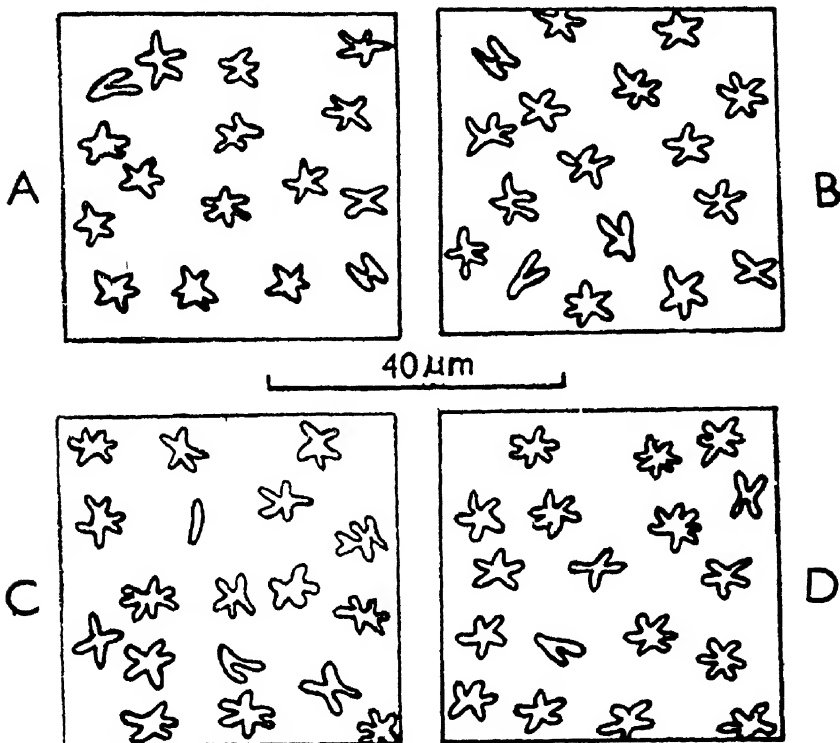


FIG 7 A-D *Heterotermes* (Rhinotermitidae, Heterotermitinae) Micrasters on dorsal surface of fore- and hindwings respectively of A, B, *Heterotermes gertrudae*, C, D, *Heterotermes indicola*

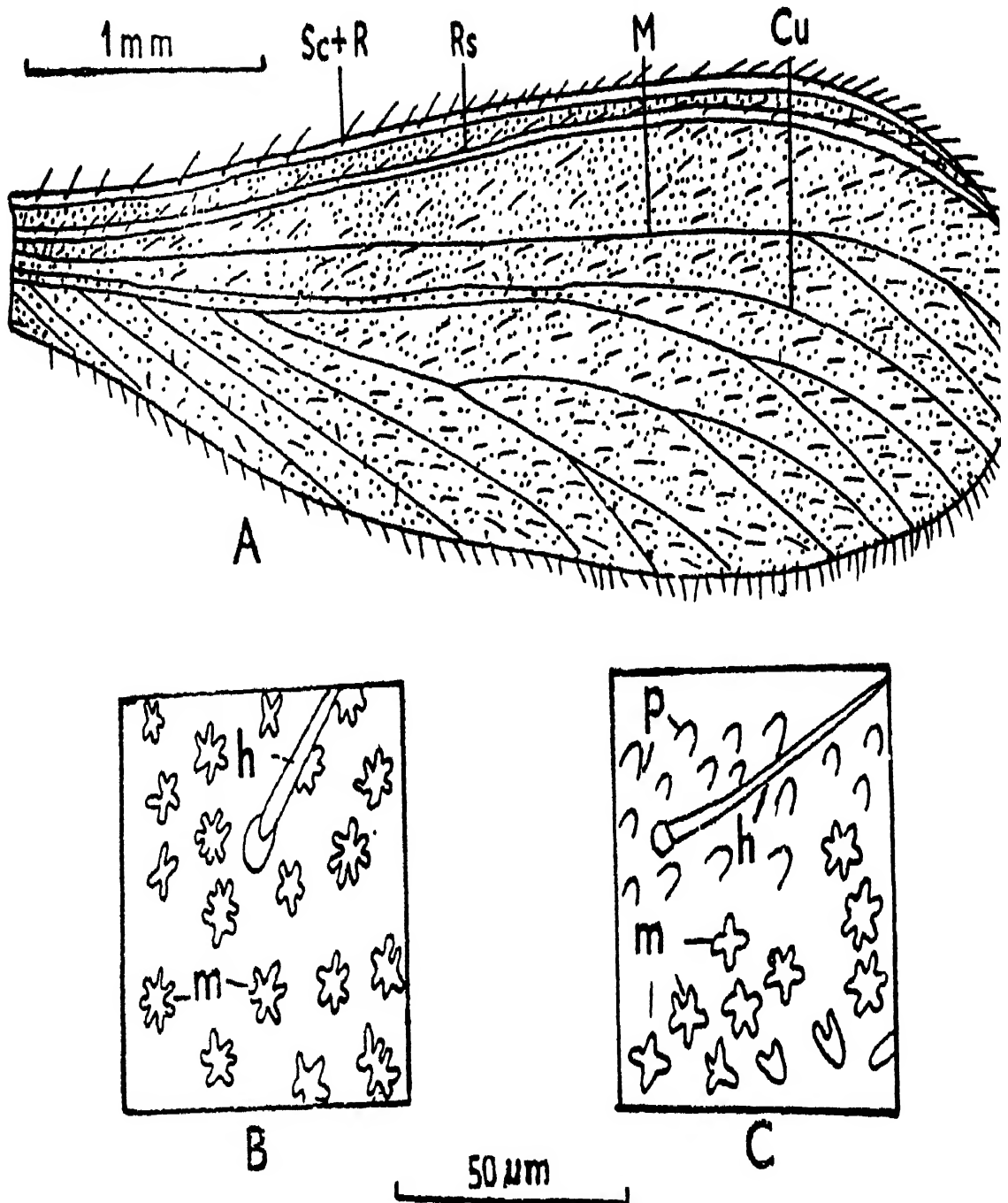


FIG 8 A-C *Termitogeton planus* (Rhinoitermitidae, Termitogetoninae) Forewing and its proximal parts, to show hairs and microsculpturing; A, Wing (without basal scale), B, Middle region; C, Anterior region (Roonwal et al 1979b)

Cu, cubitus; h, hairs; m, micrasters; M, media; p, papillae; Rs, radial sector; Sc+R, subcosta+radius

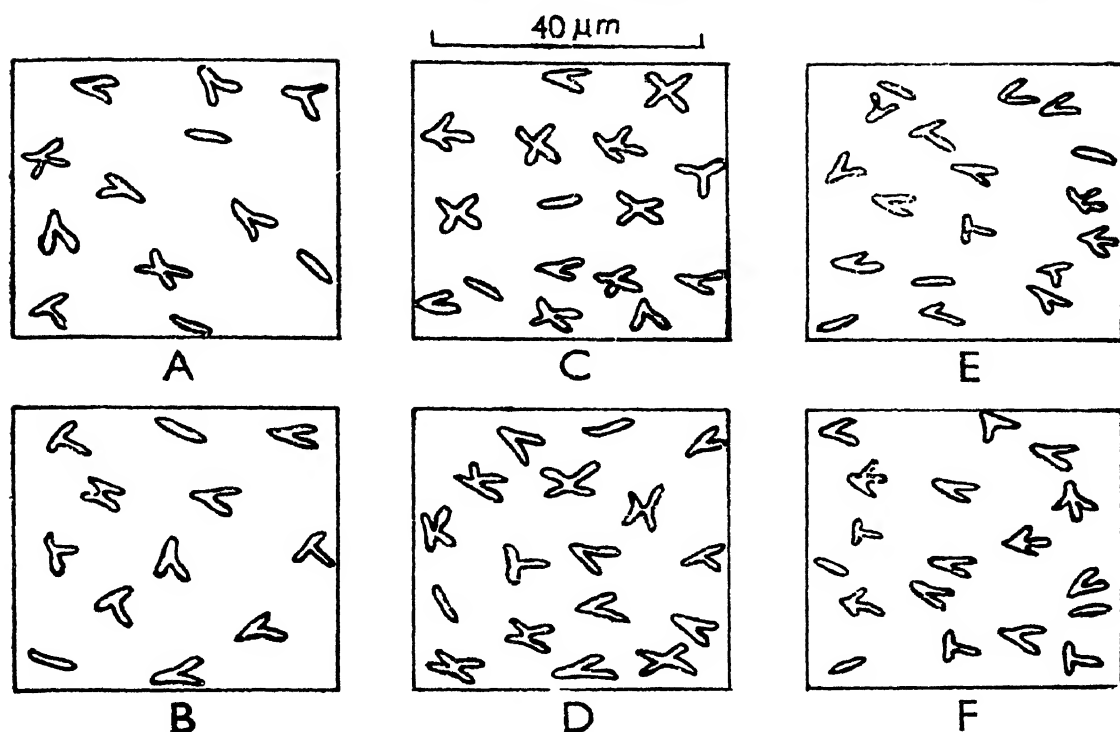


FIG 9 A-F Micrasters on dorsal surface of fore- and hindwings respectively of some Amitermitinae (Termitidae), A, B, *Alyscotermes kulandjaricus*, C, D, *Anoplotermes brevipilus*; E, F, *Anoplotermes pacificus* (Ex Roonwal & Rathore 1977)

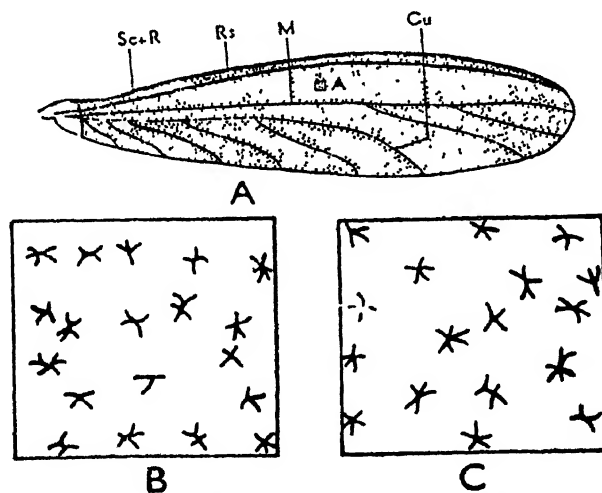


FIG 10 A-C Wing microsculpturing in *Eremotermes* (Termitidae, Amitermitinae); A, Forewing of *E. fletcheri*, to show venation and micrasters (stippling). (Small square, A, is the region from which the more magnified figures were made); B, *E. fletcheri*, forewing; C, *E. paradoxalis*, forewing. Cu, cubitus; M, media; Rs, radial sector; Sc+R, subcosta+radius

| FAMILY     | SUBFAMILY     | GENUS           | NON - ASTEROID |    |     |    |    |    | ASTEROID |      |    |   |
|------------|---------------|-----------------|----------------|----|-----|----|----|----|----------|------|----|---|
|            |               |                 | I              | II | III | IV | V  | VI | VII      | VIII | IX | X |
| TERMITIDAE | AMITERMITINAE | ALYSCOTERMES    | I              | V  | YT  | XV |    |    |          |      |    |   |
|            |               | ANOPLOTERMES    | I              | V  | YT  | XV |    |    |          |      |    |   |
|            |               | SPECULITERMES   | I              | V  | V   | XV |    |    |          |      |    |   |
|            |               | EREMOTERMES     |                |    | YT  | XV | YV | XV | X        | X    |    |   |
|            |               | MICROCEROTERMES | I              | V  | YT  | XV | XV | XV | X        | X    | X  | X |
|            | T             | ANGULITERMES    | I              | V  | YT  | X  | XV | XV | X        | X    | X  | X |

FIG 11     Distribution of various categories and types of micrasters    on the wings of some Amitermitinae and Termitinae (T)

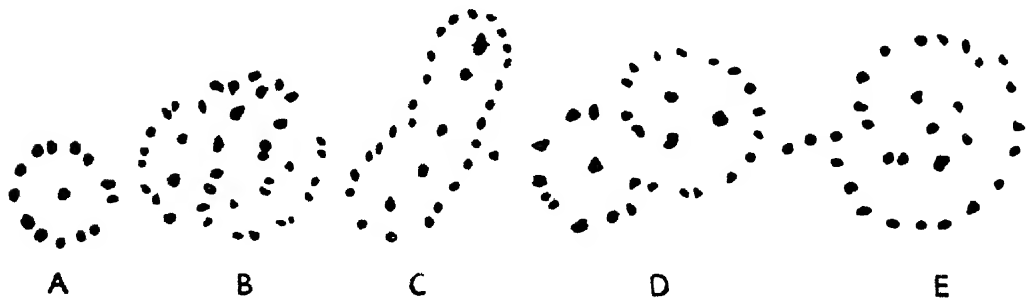


FIG 12     *Speculitermes silvestri* (Termitidae, Amitermitinae) Some types of patterns of micrasters on hindwing They tend to be curved (circles, ovoids, spirals etc , often with one or more micrasters enclosed) (Ex Roonwal & Rathore 1977)

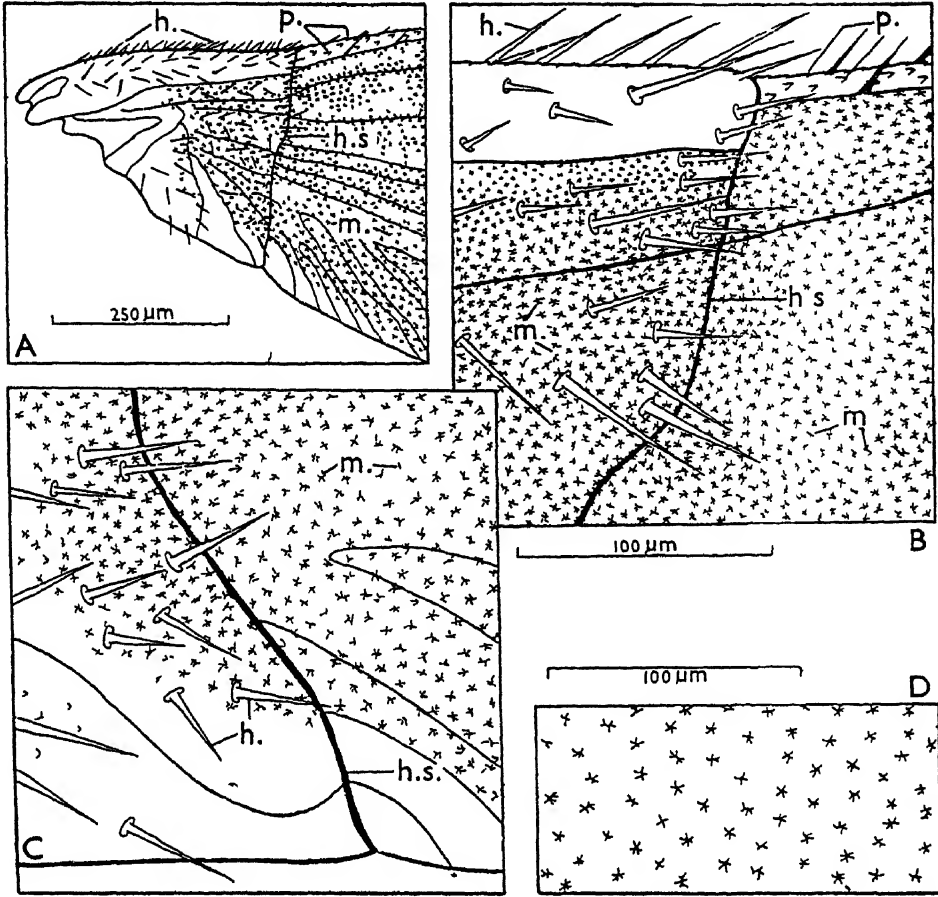


FIG 13 A-D *Eremitermes paradoxalis* (Termitidae, Amitermitinae), basal (proximal) part of right forewing, to show clumping and distortion of microsculpturing elements on wing scale and adjacent area, A, scale and part of membrane, B, Same, upper portion, more magnified, C, Lower portion, more magnified, D, Middle of membrane (normal micrasters) (Ex Roonwal & Rathore 1982)  
*h*, hairs, *hs*, humeral suture, *m*, micrasters, *p*, papillae

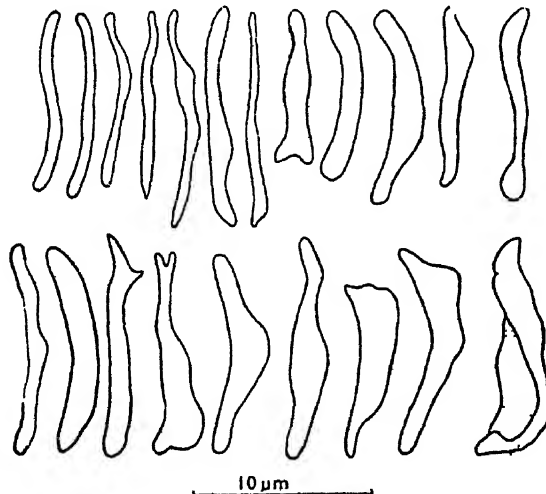


FIG 14 *Odontotermes obesus* (Macrotermitinae, Termitidae) Variation in cuticular rods on wing surfaces (Outlines from scanning electron micrographs)

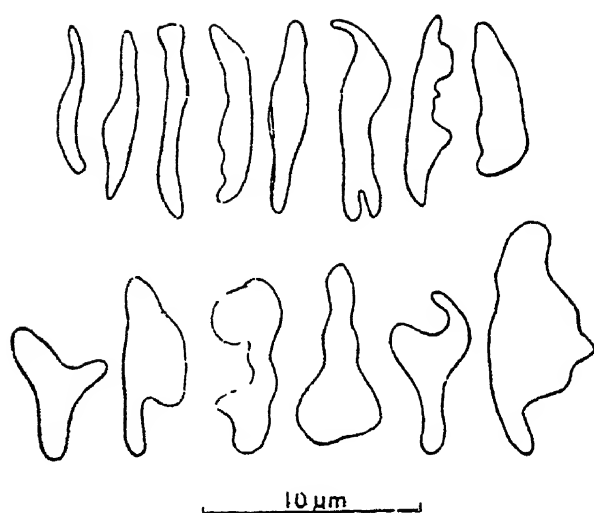


FIG 15 *Odontotermes assmuthi* (Macrotermitinae, Termitidae) Variation in cuticular rods on wing surfaces (Outlines from scanning electron micrographs)

## DISCUSSION AND CONCLUSIONS

### (A) *Evolution and Phylogeny*

The nine families of living termites can be arranged broadly in the following evolutionary order:

#### A. *Primitive families*

1. Mastotermitidae
2. Termopsidae
3. Hodotermitidae
4. Kalotermitidae

#### B. *Intermediate families*

5. Stylotermitidae
6. Serritermitidae
7. Rhinotermitidae

#### C. *Advanced families*

8. Indotermitidae
9. Termitidae

When we examine the distribution of the various types of microsculpturings, it is possible to divide the order Isoptera into three broad groups, A, B, C (figure 3) as given below, though it is not possible

to draw sharp lines of division in the intermediate families (Roonwal 1983b):

*Group A* (Primitive Group): Wings with papillae, pimples, tubercles and arrowheads. No spearheads, micrasters, microsetae and rods. Families mastotermitidae, Termopsidae, Kalotermitidae, Hodotermitidae (in part), Stylotermitidae, Serritermitidae and Rhinotermitidae (in part).

*Group B* (Intermediate Group or Higher Group I of Roonwal): Wings generally with micrasters (rarely absent) or spearheads (rare), but without rods and microsetae (these occasionally present). Families Hodotermitidae (in part), Rhinotermitidae (in part) and Termitidae (in part, especially the primitive subfamily Amitermitinae, and also Termitiane in part, Nasutitermitinae in part and Macrotermitinae in part). The tiny, single-genus family Indotermitidae (*Indo termes*), also probably belongs here, but nothing is known about its microsculpturing.

*Group C* (Advanced Group or Higher Group II of Roonwal): Wings without micrasters (rarely, simple ones present), and with or without microsetae and rods; spearheads occasionally present. Family Termitidae in part (subfamilies Nasutitermitinae in part and Macrotermitinae in part)

Thus, leaving aside the papillae (which occur universally in all termites), the remaining seven types of microsculpturings exhibit the following evolutionary sequence:

(a) Pimples, tubercles and arrowheads characterise the primitive groups. Micrasters characterise some of the more advanced among the intermediate groups and the more primitive members of the higher groups; and microsetae and rods the highest groups. Spearheads occur occasionally in the intermediate as well as the higher groups.

(b) While this broad evolutionary picture is clear enough, it cannot be said that a straight line evolutionary trend is invariably present. There are odd and puzzling exceptions which suggest that at least some of these structures may have arisen *sui generis* more than once. On the whole, however, five evolutionary types of microsculpturing elements can be broadly recognised as follows:

1. Papillae-group (distally directed):

- (a) Simple papillae (various types, e.g., finger-shaped and nipple-shaped; and pointed or thorny).
- (b) Arrowheads

- (c) Tubercles (subcrescentic or angular)
- (d) Spearheads.
- 2. Pimpules.
- 3. Micrasters:
  - (a) Simple (nonasteroids with none to a few arms). This includes the new type (discovered by scanning electron microscopic examinations) with a basal leafy or flowery vase and 0-4 central arms or rods protruding from the centre of the vase.
  - (b) Complex (asteroids with many arms, up to 8).
- 4. Microsetae
- 5. Rods (thin or thick)
  - (a) Subvertical (vertical-oblique)
  - (b) Subhorizontal (horizontal-oblique)

(B) *Comparison with Allied Insect Orders* (figures 16-19)

It is revealing to compare microsculpturing in the Isoptera with the condition occurring in insect orders which are closely allied to it, even though the information available is not as detailed and extensive as in the Isoptera. We have information on three orders, and the fact which stands out is that termites are unique in the richness (both in density and variety) of the microsculpturing elements which, even when present, are much poorer in number and far simpler in variety in the other orders.

*Dictyoptera* (figure 16): This includes the cockroaches (Blattaria) which are widely believed to be ancestral to termites. A single species, the Small Indian House Cockroach, *Supella longipalpa* (Fabricius) (family Blattidae), has been studied (Roonwal & Rathore 1983). Besides hairs, microsculpturing here consists of seven types of structures, viz., papillae, pimpules, rods, microsetae, polygons, parallel ridges and an irregular reticulum, the last three elements being rather diffuse and ill-defined. Their density is low and never exceeds ca. 1000/mm<sup>2</sup>. The microsculpturing thus is less discrete and of poor quality.

*Zoraptera* (figure 17): This is a small but widely distributed order with a single genus, *Zorotypus* (family Zorotypidae), and only about 22 species. The order is closely allied to termites, and, like the latter, its members shed their wings at a basal suture. A single species, *Zorotypus hubbardi* Caudell (from Hunter, Virginia, USA) has been studied



(Roonwal 1983b). Both surfaces of wings are covered all over with numerous small hairs (length 15-60  $\mu\text{m}$ , density ca. 1030/ $\text{mm}^2$ ), and no other type is present.

*Embiopoda* (figures 18 and 19): This is a medium-sized order which is closely allied to termites. Several species belonging to two families and three genera have been studied: Embiididae (genera *Parembia* and *Embia*) and Oligotomidae (genus *Oligotoma*) (Roonwal & Rathore 1984). Microsculpturing here consists (besides hairs or macrotrichia, 20-130  $\mu\text{m}$  long) of two other types of structures, viz., microtrichia (small, fixed hairs, length 8-23  $\mu\text{m}$  and macroscimitars (short to long, curved, scimitar-like structures with a swollen base; length 16-150  $\mu\text{m}$ ). The last mentioned structures are present in one or two rows at the anterior wing margin, while the other two are present all over the membrane on both sides (density of macrotrichia ca. 88-300/ $\text{mm}^2$ , of microtrichia ca. 1016-4050/ $\text{mm}^2$ ). No other type is present, and the macroscimitars constitute a type which seems to be unique in insects.

### (C) Biological and Ecological Consequences

We may now consider whether this dense carpeting of microsculpturing elements on termite wings subserves any useful function. If it does not, then how has this elaborate patterning, involving a million or more elements on a single wing, evolved by natural selection? I have no clear-cut answers, but will try to assess the consequences as far as possible.

It is possible, and indeed very probable, that the occurrence of these innumerable bodies on the wing surface adds some strength to them (in the same way as the veins do) and prevents the wings from being torn asunder in high wind. The heavily weighted wings may also prevent termites from being blown away. On the other hand, microsculpturing seems to result also in a number of apparent disadvantages.

No study of the aerodynamics of termite wings has been made. But field observations show that wing flapping in termites is slow. Flight is generally weak and erratic and probably depends largely upon air currents. (Termite flights occur periodically, usually once or twice a year, in swarms, depending upon the production of winged images or alates in a colony. For details of swarming characteristics, especially in deserts, see Roonwal 1976, 1979, 1983c).

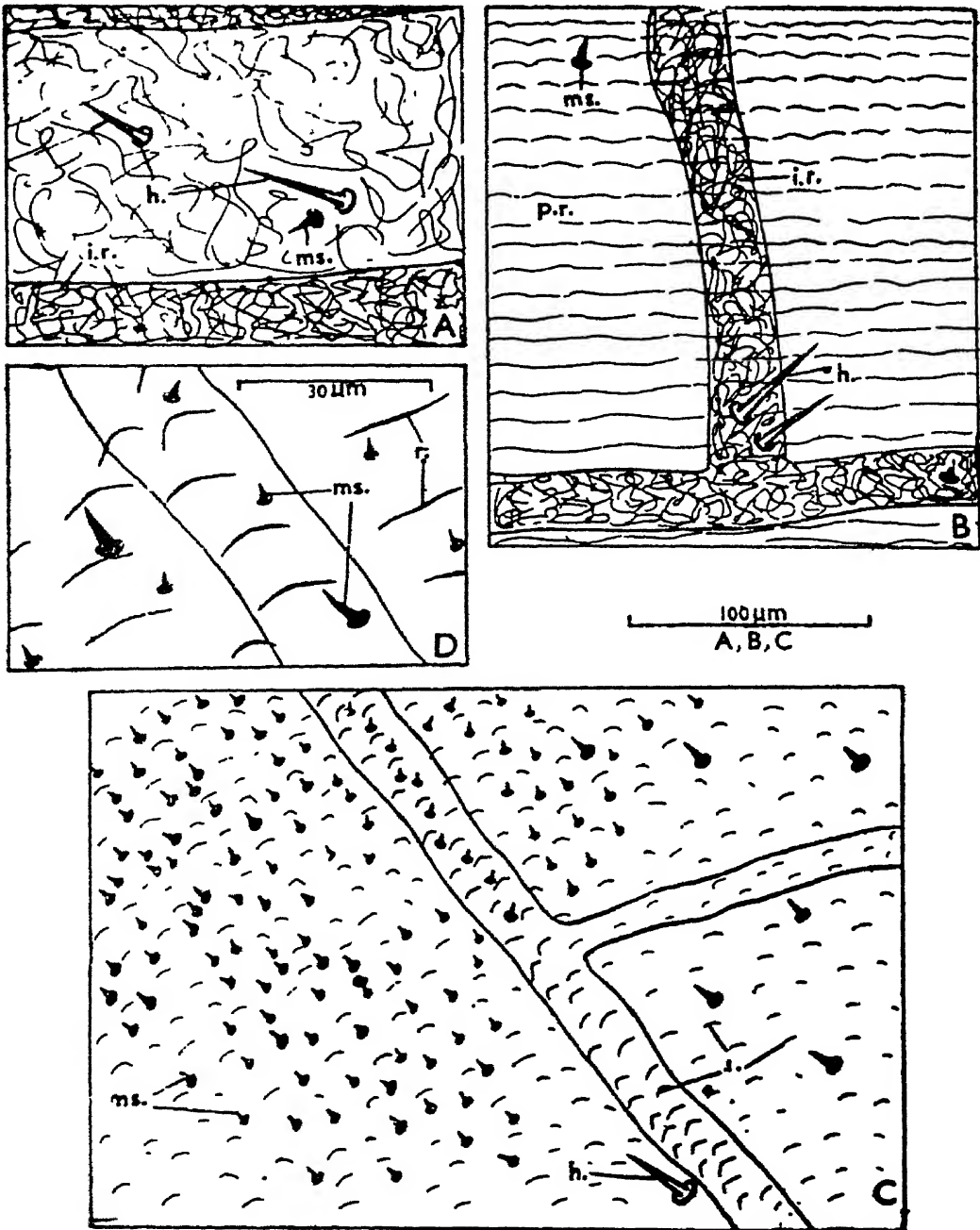


FIG 16 A-D *Supella longipalpa* (Diptera, Blatidae). Parts of dorsal surface of male wing, to show microsculpturing. A, Near anterior end; B, Near distal end; C, Middle of wing; D, Same, a part more magnified. (Ex Roonwal & Rathore 1983)

*h*, hairs, *ir*, irregular reticulum; *ms*, microsetae; *pr*, parallel ridges, *r*, rods

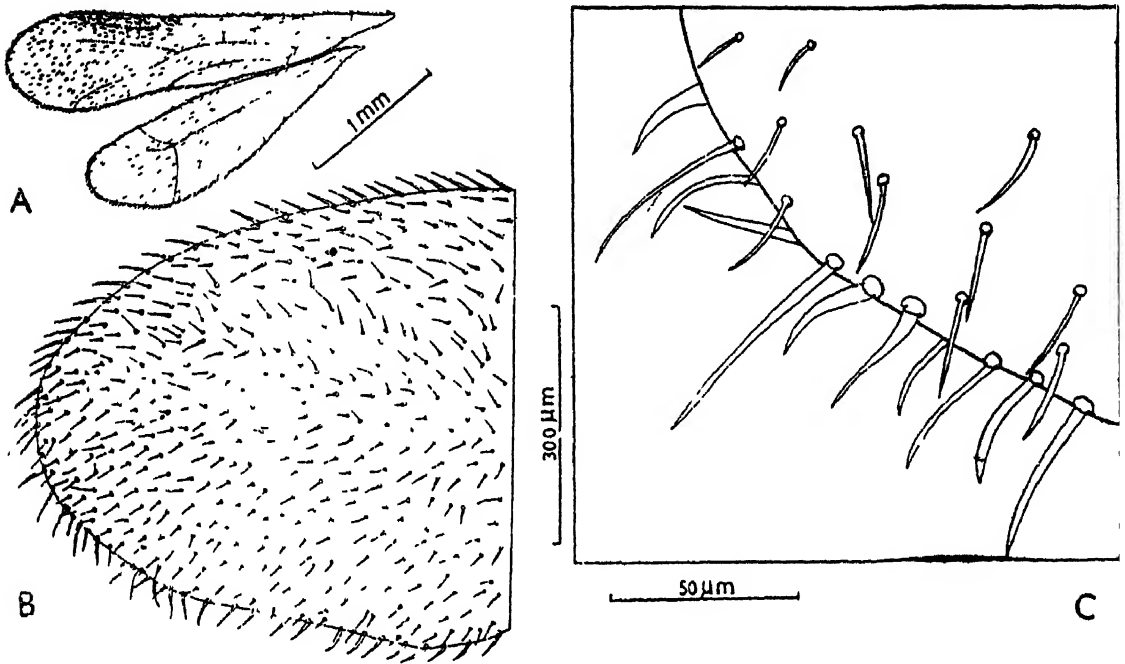


FIG 17 A-C *Zorotypus hubbardi* (Zoraptera), Left wings, to show dense covering of hairs, A, Both wings, B, Distal part of forewing, more magnified; C, Same, a small portion near margin, to show the long and short hairs (Ex Roonwal 1983b)

The occurrence of microsculptures results in two obvious consequences: (i) It adds enormous dead weight to the otherwise light and thin wings; and (ii) it causes extreme rugosity of wing surface (figure 20) and restricts and modifies air flow on the otherwise almost smooth wings, i.e., it increases drag. Both these consequences must greatly impede the aerodynamic efficiency of wings, but an understanding of the precise manner in which the impediments may occur must await further research. It is highly likely that the obstacles result in slowing down flight, a conclusion borne out by field observations.

These effects have important biological consequences:

(i) They prevent the dispersal of termites by purely intrinsic means to distant and perhaps hostile ecological environments, and swarms can start fresh colonies only within a short distance of the swarming sites, perhaps no more than a few kilometres. Any dispersal to distant places (and several such cases are known), occurs by extrinsic agencies alone, such as strong wind, driftwood and commerce.

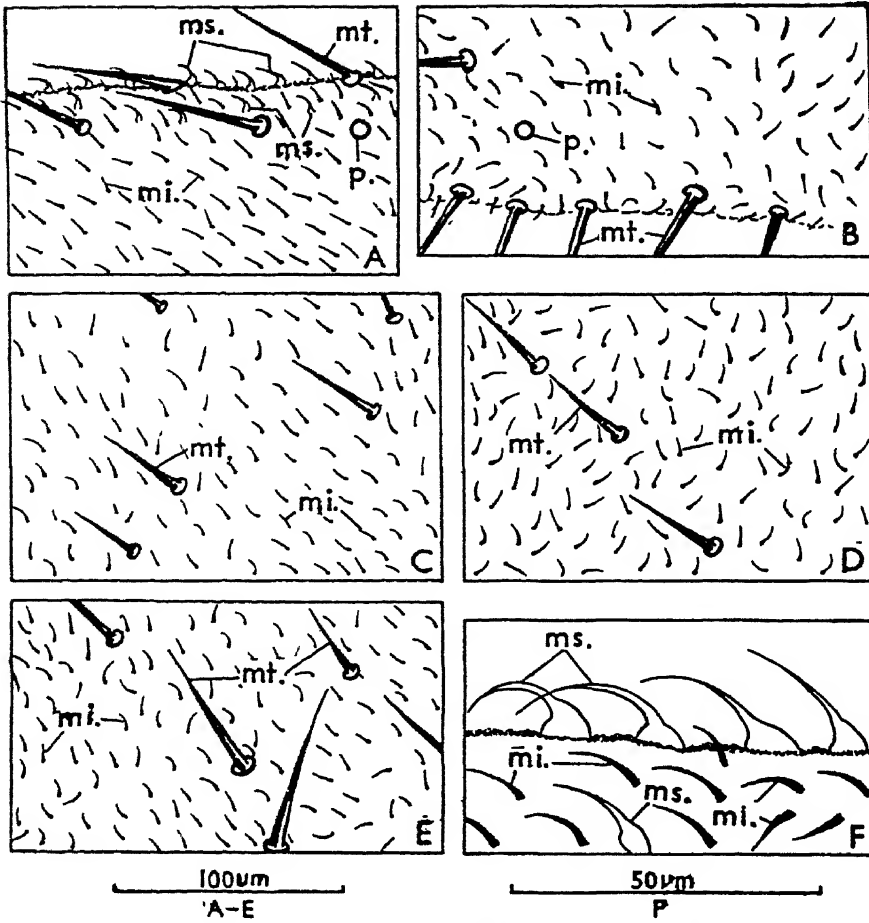


FIG 18 A-F *Parembia valida* (Embioptera) Views of different parts of dorsal surface of forewing, to show microsculpturing (Ex Roonwal & Rathore 1984)  
*mi.*, microtrichia, *ms.*, macroscimitars, *mt.*, macrotrichia, *p.*, basal platelet of a macrotrichium (the latter removed)

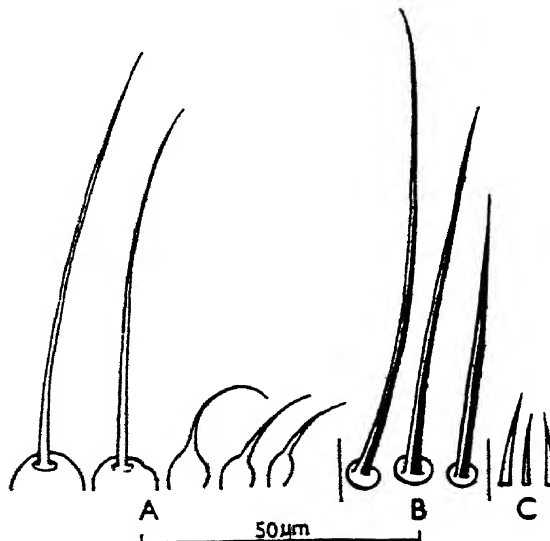


FIG 19 A-C Types of wing microsculpturing in Embioptera. A, macroscimitars, B, Macrotrichia (hairs), C, Microtrichia (ex Roonwal & Rathore 1984)

(ii) The slow, erratic flight makes the termites easy prey to predators on the wing, e.g., birds and bats; and on falling to the ground they become easy prey to lizards, squirrels and other mammals, including man. Thus, perhaps 80 per cent or more of the swarming individuals are eliminated.

(iii) The primitive families have fewer and smaller microstructures, e.g., papillae only, or papillae and pimples, etc. For this reason, they, theoretically, have faster flights, and, thus, better powers of dispersal.

The study of microsculptures has a practical significance in taxonomy, and these structures often provide excellent morphological characters for differentiating species and genera (Roonwal et al. 1974, Roonwal 1983b). Species of the same genus can be distinguished by differing degrees of microsculpturing. Thus, *Microcerotermes cavus* has only three simple types of micrasters (types III to V), while *M. nicobaricus* has a great variety (types I to X). In *Heterotermes gertrudae* the papillae are finger-shaped, while in the closely allied species, *H. Indicola*, they are pointed and thorny.

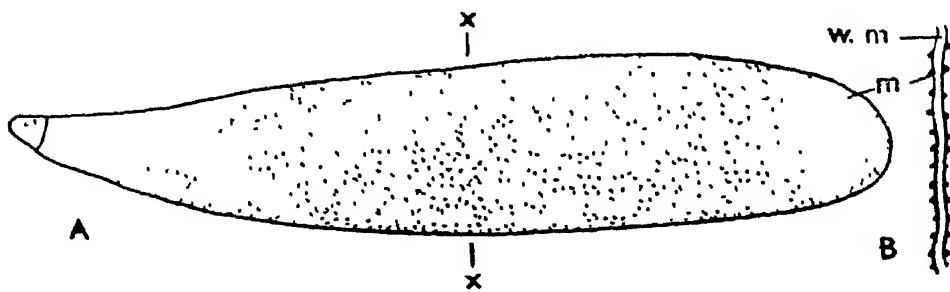


FIG 20 A-B *Angulitermes jodhpurensis* (Termitinae, Termitidae), forewing A, Outline of wing, showing the distribution of micrasters (suppling) all over wing, veins not shown Diagrammatic, B, Section of wing across line X-X in figure A, to show the micrasters on both surfaces of wing (from a photomicrograph, greatly magnified)

m, micrasters, w, wing membrane (double-layered)

In the characterisation of the termite order Isoptera, the occurrence of wing microsculpturing must now be regarded as an essential character of the order; and it has been re-defined accordingly (Roonwal 1983a, b).

Apart from wings, such cuticular micro-sculpturings (in the shape of papillae or scaly, overlapping serrations, like tiles on a roof) have been found so far on the tibial spines and apical spurs of two primitive species of termites, viz, the Himalayan termite, *Archotermopsis wroughtoni* and

the Australian termite, *Mastotermes darwiniensis* (Roonwal 1972, 1974, 1983a, b and in press, and Roonwal & Bose 1970).

### ACKNOWLEDGEMENTS

I should first pay tribute to my co-workers who have collaborated with me in these studies over a long period of years. They are: Dr O B Chhotani (Deputy Director, Zoological Survey of India, Calcutta), Dr S C Verma (of the Northern Regional Station, Zoological Survey of India, Dehra Dun), Dr N S Rathore (of the Desert Regional Station, Zoological Survey of India, Jodhpur), and Dr M L Thakur (of the Entomology Branch, Forest Research Institute, Dehra Dun).

I would also like to thank Dr P K Sen Sarma and Mr S S Bisen (of the Forest Research Institute, Dehra Dun), Dr W Wittmer (of the Naturhistorisches Museum, Basel, Switzerland), and Dr R Guggenheim and Mr G Lüönd (of the Laboratory of Scanning Microscopy of the Geological-Palaeontological Institute of the University of Basel) for their valuable assistance in scanning electron microscopic studies. To Professor K N Katiyar and Dr S M Mohnot (of the Department of Zoology, University of Jodhpur), I am indebted for assistance in photomicrography and in other respects.

### References

1. Chapman R F 1982 *The Insects. Structure and Function*. (3rd ed.). xix+919 pp. (London: Engl. Language Book Soc. and Hodder and Stoughton)
2. Emerson A E 1967 Cretaceous insects from Labrador. 3. A new genus and species of termite (Isoptera: Hodotermitidae); *Psyche*, Philadelphia, **74** 276-279
3. Grassé P P 1949 Ordre des Isopteres ou termites; in *Traité de Zoologie* ed. P P Grassé, Vol. 9, pp 408-544 (Paris: Masson et Cie)
4. Imms A D 1965 *A General Textbook of Entomology, including the Anatomy, Physiology, Development and Classification of Insects*. (9th ed., revised by O W Richards and R G Davies). x+886 pp (London: Engl. Language Book Soc. and Methuen & Co.)
5. Roonwal M L 1972 Papillae on tibial spines and spurs in the Australian termite, *Mastotermes darwiniensis*; *Sci. & Cult.*, Calcutta **38** 255
6. —1974 occurrence of papillae on tibial spines and spurs in the Australian termite, *Mastotermes darwiniensis*, with remarks on some other primitive species; *J. Zool Soc. India*, Calcutta **24** (1972) 155-160

7. —1976 Field ecology and eco-biogeography of Rajasthan termites: A study in desert environment; *Zool Jahrb. (Syst.)*, Berlin **103** 455-504
8. —1977 Micro-sculpturing on termite wings (Isoptera), and its significance, *Proc Zool Soc*, Calcutta, **28** (1975) 31-39
9. —1979 *Termite Life and Termite Control in Tropical South Asia* xii+177 pp., 8 pls. (Jodhpur: Scientific Publishers)
10. —1981 Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). XI Some hitherto unstudied genera and species in five families; *Proc Indian natn Sci Acad* **B47** 467-473
11. —1983a Universal occurrence of external cuticular microscopic papillae and allied structures as an essential character of termites (Isoptera, Insecta), and redefinition of the Order, *Zool Anz*, Jena **211** 137-144
12. —1983b Evolution and systematic significance of wing micro-sculpturing in termites XIII Order Isoptera, *Proc Indian natn Sci Acad* **B49** 359-391
13. —1983c The ecology of termite swarming in the Indian Desert; in *Insect Interrelations in Forest and Agroecosystems* ed P K Sen-Sarma, S K Sangal and S K Kulshreshtha pp 9-13 (Dehra Dun Jugal Kishore & Co)
14. —1985 Scanning electron microscopic studies of wing microsculpturing in termites I. Genera *Heterotermes* and *Microtermes*; *Proc Indian natn Sci Acad* **B51** 6-14
15. — Electron microscopic study of cuticular papillae on tibial spines and spurs of some termites (*Mastotermes* and *Archotermopsis*); *Indian J Ent*, New Delhi (In press)
16. — Scanning electron microscopic studies of wing microsculpturing in termites (Isoptera). II. Genera *Amutermes* and *Eremotermes* (Termitidae, Amitermitinae). (In press)
17. — and Bose G 1970 Scaly serrations on spines and apical spurs of tibia in the Himalayan termite, *Archotermopsis wroughtoni*, *Sci & Cult*, Calcutta **36** 673
18. — and Chhotani O B 1967 Wing microsculpturing in termite genera *Odontotermes*, *Hypotermes* and *Microtermes* (Termitidae: Macrotermitinae) and its taxonomic value; *Zool Anz.*, Jena **178** 236-262
19. —, — and Verma S C 1981 Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). X. Subfamily Nasutitermitinae of family Termitidae; *Proc Indian natn. Sci Acad.* **B47** 341-369
20. — and Rathore N S 1977 Third study of evolution and systematic significance of wing micro-sculpturing in termites. Micrasters in some Rhinotermitidae and Termitidae; *Zool. Anz.*, Jena **198** 298-312

21. — and — 1978 Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). New types in the Kalotermitidae and Hodotermitidae; *Zool Anz*, Jena **200** 219-232
22. — and — 1982 Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). XII. Sculpturing on wing scales; *Proc Indian natn Sci Acad.* **B48** 322-343
23. — and — 1983 Wing micro-sculpturing in the small house cockroach, *Supella longipalpa* (Dictyoptera, Blattidae); *Proc. Indian Acad Sci. (Anim. Sci.)*, Bangalore **92** 333-342
24. — and — 1984 Wing micro-sculpturing in Embioptera (Insecta), with a new type of structure and ecological and taxonomic considerations, *Indian J Ent* **45** (1983) 323-334
25. — and — Wing microsculpturing in the Brazilian termite family Serritermitidae (*Serritermes serrifer*, Isoptera) and its bearing on phylogeny (in preparation)
26. — and Sen-Sarma P K 1960 in *Contribution to the Systematics of Oriental Termites*. xiv+407 pp., 65 pls. (New Delhi: Entom. Monogr. No. 1, Indian Counc. Agric. Res., Publ. by Manager of Publ., Govt. of India, Delhi)
27. — and Verma S C 1976 Imago of the wood termite *Heterotermes gertrudae* Roonwal (Rhinotermitidae: Heterotermitinae) from North-West Himalayas; *Rec. zool Surv India*, Delhi **69** 241-247
28. — and — 1980a Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). VIII. Subfamily Amitermitinae of family Termitidae, *Proc Indian natn. Sci Acad.* **B46** 250-263
29. — and — 1980b Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). IX. Subfamily Termitinae of family Termitidae; *Proc Indian natn Sci Acad* **B46** 455-469
30. — and — 1983 New data on wing microsculpturing in termites (Kalotermitidae, Rhinotermitidae and Termitidae); *Annals Ent*, Dehra Dun **1** 27-34
31. — and — Scanning electron microscopic studies of wing microsculpturing in termites (Isoptera). III. Genera *Odontotermes* and *Macrotermes* (Termitidae, macrotermitinae) (in preparation)
32. — —, and Rathore N S 1974 On a new systematic character in termites, the micrasters; *Z. zool. Syst. u Evolutionsforsch.*, Berlin **12** 55-76
33. —, — and — 1980 Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). VII. Subfamily Macrotermitinae of family Termitidae; *Proc Indian natn. Sci.Acad.*, **B46** 149-163
34. —, — and Thakur M L 1979a Evolution and systematic significance of wing microsculpturing in termites (Isoptera). V. Families Mastotermitidae,



Termopsidae, Hodotermitidae and Stylotermitidae; *Proc. Indian natn Sci Acad.* **B45** 115-128

35. —, — and — 1979b Evolution and systematic significance of wing micro-sculpturing in termites (Isoptera). VI. Family Rhinotermitidae; *Proc. Indian natn. Sci Acad.* **B45** 332-353
36. Weesner E M 1969 External anatomy; in *Biology of Termites* ed K Krishna and F M Weesner Vol. 1, pp 19-47 (New York & London: Academic Press)
37. Weidner H 1970 Isoptera (Termiten); in *Handb. d. Zool*, 4 (Sec. 2, Pt. 2), (No.14), pp 1-147 (Berlin: Gruyter & Co)



**Ishwar Prakash** (b. 17 December 1931) did D.Sc.(1983) from University of Rajasthan, Jaipur. He is INSA Senior Scientist (1992), and was formerly Professor of Eminence (1980-91), Coordinator and Principal Animal Ecologist (1977-80) at Central Arid Zone Research Institute, Jodhpur.

Prakash has carried out pioneering research on the ecology of desert vertebrates, environmental analysis and desertification. His researches have ranged from faunistic studies to biochemical communication among desert rodents, physiology of desert survival and wildlife conservation to rodent pest management

He is considered to be an authority on desert ecology. He has authored/edited 17 books and monographs

Prakash is Fellow of National Academy of Sciences (India) and National Academy of Agricultural Sciences (India); Foreign Fellow, International Thermological Congress, Russian Academy of Sciences. He is the recipient of Rafi Ahmed Kidwai Memorial Prize (ICAR) (1975); Har Swaroop Memorial Lecture Award (INSA) (1991).

*Ishwar Prakash was elected to the fellowship of the Academy in 1982.*

**Address:** INSA Senior Scientist, Zoological Survey of India, Paota B Road, Jodhpur 342 010, Tel. (0291) 45144; Res. Vidya Kailash, Near Sati Mata Temple, Ratanada, Jodhpur 342 001; Tel. (0291) 31654.

## **THE CHANGING ECOLOGY OF THE THAR DESERT**

ISHWAR PRAKASH

The Great Indian or the Thar desert situated in northwest India, encompassing parts of Gujarat, Rajasthan, Haryana and Punjab States, used to be well watered country during geologically recent times. Due to geotectonic activities and changes in the monsoon system, the water table sank and gradually the rivers vanished. The ancient civilisation of Mohenjodaro, Harappa and Kalibanga and their prosperity also indicate that quite different type of physical conditions prevailed in the region. However, with the availability of more information on the physical and biological evidences about the genesis of Thar desert, it is now clear that this desert is not 'man made' as is often claimed but has been created due to multiple factors and that the desert conditions prevailed in this region much earlier than 10,000 years (Gupta and Prakash, 1975).

The Indian desert is not a continuous stretch of sand but its topography is interspersed by hillocks, pediment zones and gravel plains. The dry beds of ephemeral rivers constitute typical habitat, full of soil moisture during monsoon and totally dry for 9 months in the year. A vast variation in habitats, sometimes in very close geographical proximity, create conditions suitable for isolation, thus accelerating the process of speciation. The major habitats found in the Thar are: Hilly outcrops, Sand dunes, Sandy plain, Gravel plain, Saline patches/Lakes, and Ruderal complex.

All these major formations hold typical flora and fauna in different rainfall zones. Monsoon is a mighty event for the desert biota. With the first shower of the season, the forbidden sea of sand assumes a gay cloak of most magnificent greenery. The sweet patter of the long-awaited rain drops break the slumber of the dormant seeds lying buried under the sand and the desert is transformed into a garden as if by the touch of a sorcerer's magic band. The tiny toads appear from nowhere, the eggs of insects and snails come to life. Almost all vertebrate females enter into oestrous and a fast and time-bound reproduction activity starts (Prakash 1960).

The lush monsoon vegetation of the desert is transient (Saxena and Prakash 1992). The desert scene starts changing soon after monsoon, after a second summer in October, the greenery fades and the Thar starts

looking brown, and by winter bare sandy patches appear. With the advent of spring emerges a fresh crop of sprouts but these are quickly grazed over by the wandering, ever-increasing herds of hungry livestock and the all destructive hordes of rodents. The desert soon conforms to its popular image of an area of bare, rolling sand dunes, undulating sandy plains, dry shrubs, a few thorny trees and the seductive mirage. The wind grows strong during May and June with the afternoons often darkened by the howling storms and the world seems lost under a blanket of dust - the full fury of the desert is now let loose on man and beast alike (Prakash 1977).

Contrary to popular imagination, the desert is an animal lover's paradise. One only has to look around to wonder at the rich and variegated animal life that this apparently inhospitable tract harbours within its orbit. At day's end, when the sun exhausts its fury and the howling winds calm down to a soothing breeze, the desert really comes to life although day-active forms are not hard to come by. The denizens of the desert have learnt quite a few tricks, in the course of their evolutionary history, to ekeout a successful living in the terrain of their choice. These include adaptations of body forms, behaviour and the functioning of internal organs. Nocturnalism, subterranean life, omnivory are some of the survival mechanisms of animals for withstanding the vagaries of desert environment. Behaviourally, their circadian rhythm is switched over to combat the heat as well as cold. Typical desert animals thrive without drinking free water. They produce metabolic water through hydrolysis of fats which they accumulate during the previous winter and conserve it for utilisation during the summer (Prakash 1986).

The biota found in the desert is an admixture of Ethiopian, Saharan, Oriental and Peninsular elements (Prakash 1963). The analysis of faunistic elements reveal that 64 per cent biota bear saharan affinities. The endemic element is also fairly large which indicates that the Thar desert is not a young desert (Prakash 1974).

In the past, the lion, cheetah, caracal, wild ass were fairly common in the arid terrain. Trigger-happy man has played havoc with his wildlife heritage and the gory drama has, in a way, been helped by certain ecological changes in the land use pattern in the desert region. The striking increase in the human population from 3 to 16 million during the last 90 years has transformed the grassland into a cropland, leaving very little grazing land for wild species. Paradoxically, along with the reduction in the grazing area, the livestock population has registered an alarming

increase, from 10 to 23 million in past 30 years resulting into drastic depletion of vegetation resources in certain habitats, the natural succession trends have been reversed. Consequently, population numbers of unwanted animal species like field rodents and termites have substantially increased. These pests further deplete the vegetation resources.

The Thar desert is by far the most populated desert of the world. With its human population of 75 per km<sup>2</sup>, man is bound to interfere with the desert ecosystem in a big way. In fact, this desert is being exploited as a future land bank. As a consequence, it is under serious biotic pressure on its natural resources. In this context a ray of hope has emerged due to irrigation of a large region in the northwestern Thar desert through the Indira Gandhi Nahar (IGN) which is fast transforming a vast grazing land into a cultivated ecosystem. The impact of availability of irrigation water is apparent through the replacement of the xeric plants and animals by mesic elements. Whereas the productivity of land and the crop production has enhanced manifolds, its harmful repercussions are now clearly visible (Mann and Prakash 1983).

#### THREAT OF SEWAN GRASSLAND DEPLETION

Extensive irrigation is now possible due to availability of very good quality of water from the Indira Gandhi Canal which envisages to transform about 11% of the barren un-inhabited area of Western Rajasthan into a vast greenery. For bringing this water to the desert, most of the 204 km of Rajasthan feeder and 445 km of Indira Gandhi Main Canal and about 8000 km long branches and tributaries have been constructed. These will command over 2 million ha of which 1.3 million ha is culturable.

Vast grassland comprising of a single species, *Lasiurus indicus* (*Sewan*) occur in the 100 mm rainfall zones of Bikaner, Jaisalmer and Barmer districts. In the same area the IG canal irrigation has enhanced soil moisture regime and this tussocky and highly nutritive grassland is fast vanishing. If this is allowed to happen, it will be a severe blow to the genetic diversity since this perennial grass grows only in this region (Prakash 1992). Moreover, since drinking water was not available in the erstwhile IGN region, it endured almost no pressure from human and livestock populations. Now that drinking water is available, thousands of animals migrate to this region during summers and drought periods. The onslaught of grazing animals is of the magnitude that this Sewan grassland will vanish sooner or later. It is, therefore, suggested that:

- a) Water from the IG canal should be pumped to areas of existing Sewan grassland and where no cropping is envisaged. A minimal irrigation by sprinklers so as to keep this perennial grass green for the whole year. Irrigation requirement for Sewan would be about 10% of what is required for rabi crop production. The Sewan grasslands, so greened, should not be allowed to be grazed. Grass should be clipped at the most nutritive stage and provided outside the Sewan area. This suggestion invites lot of Governmental action and is likely to be negated even at the initial stage because it requires strenuous work, high investment and stay of Government officials in remote areas (Prakash 1989).
- b) Vast areas in the I.G. Canal command zone are unsuitable for cropping, chiefly because soil depth is not adequate or the murram calcareous layer is too close to the soil surface and irrigating such soils zones will initiate water logging and salinity problems even in the first year. Such lands should be utilised for silvipastoral use along with plantation of top feed trees. For this purpose only drip irrigation should be utilised. No flood-irrigation should be allowed to prevent water logging and salinity problems. In the inter-row zones of trees, Sewan should be planted. This irrigated Sewan grassland is expected to produce so much of highly palatable feed that it would not only satiate the requirements of grazing animals in Rajasthan but can be stored in Fodder Banks to combat the drought situations. It is high time that an urgent "thinking" is given to livestock production with an ecological approach and a practical bias. The opportunities should not be lost (Prakash 1992).

### EXTINCTION OF WILDLIFE

The rangelands which are being transformed into a huge irrigational zone harbour specialised form of wildlife species, highly adapted to successfully survive in the extreme xeric environment. Most important amongst them are the desert cat, *Felis libyca ornata*; desert fox, *Vulpes v. pusilla*; the caracal, *Felis caracal*; the Indian gazelle, *Gazella bennetti*, the great Indian Bustard, *Adreotis nigriceps*. Besides a large number of reptiles and migratory birds. With the changing floristic composition, human presence, and edaphic characteristics; these animals will either migrate out of the affected area or will perish (Prakash and Ghosh, 1980).

To combat this situation, the best solution is to capture adequate numbers and re-introduce them in areas where they were commonly found in the recent past (Prakash 1988). Prior to materialise and implement this idea, however, it will be essential to evaluate the

- food and shelter resource in the new habitat
- carrying capacity of land vis-a-vis grazing pressure of livestock
- predator population
- standardization of capturing/tranquilising techniques
- mode of transport

The process of re-introduction should be implemented at a slow pace. Only a few herds should be translocated and their ability to adjust to new environment should be carefully watched and studied. If this pilot experiment succeeds and the animals start breeding in the newer area, more animals should be introduced.

In addition to re-introduction of threatened animals in other areas, a few National Parks can be established in IG region, in lands which are not suited for introducing irrigation, for *in situ* conservation of the typical biodiversity of the desert ecosystem. The desert National Park (Desert Biosphere Reserve) situated in the Jaisalmer and Barmer districts will not serve this purpose.

### RISK OF EPIDEMICS

The typical desert rodents (*Gerbillus*, *Meriones*, *Tatera*) are being replaced by mesic elements (*Millardia*, *Bandicota*). The gerbils *Meriones hurrianae* and *Tatera indica* are the sylvan vectors of the plague bacillus, *Yersinia pestis* whereas *Millardia melitana* and *Bandicota bengalensis* are highly susceptible to this bacillus (Prakash 1988a). The same habitat is being occupied by them, one resistant and the other highly positive, there is a high risk of spread of dreaded plague even in human population. Excellent cost-worthy techniques for rodent pest management are now available (Prakash and Ghosh 1992; Prakash and Mathur 1987). There is an urgent need for taking up extensive rodent control campaigns and to educate the masses to ensure their participation.

It is inevitable that there will be a biological invasion of the mesic elements into the altered desert ecosystem. Sri Ganganagar region, where the replacement of species is fairly at an advanced stage, has opened new

doors for study. Cutaneous leishmaniasis is on the increase in the rural population. Hitherto unknown and new pests, weeds and plant diseases are erupting in the region. A continuing monitoring of the impact of the ecological changes is essential so that remedial measures can be taken up.

Governments are alert about the changing situations in the desert, and have established a Desert National Park in Jaisalmer and Barmer districts. Department of Environment and Forest, Government of India have proposed to upgrade this Desert National Park into a Desert Biosphere Reserve extending over 3100 Km<sup>2</sup> (Anonymous 1988). Out of this, the core area of 600 km<sup>2</sup> will be rigidly protected from all bio interferences. It is envisaged that a fairly large number of desert adapted species will be able to rehabilitate in this protected area. More important are the Great Indian Bustard, Common sandgrouse, grey partridge, gazelle, desert fox, desert cat, caracal and a number of reptilian species.

It is inevitable that with the present growth of human and livestock population, the face of the desert, as we know it today, will change sooner or later. However, if the typical bio-diversity occurring in the desert could be preserved in the Biosphere Reserve, we would be able to keep a part of this heritage with us for ever and for the benefit of future generations to admire and enjoy.

## References

1. Anonymous 1988 Thar Desert Biosphere Reserve pp. 1-100 (New Delhi: Ministry of Environment and Forests).
2. Gupta, R K and Prakash I 1975 *Environmental Analysis of the Thar desert* pp 1-628 (Dehradun: English Book Depot)
3. Mann H S and Prakash I 1983 *Halting the March, Eco-development in the Thar* pp. 1-36 (New Delhi: Department of Environment and WWF)
4. Prakash I 1960 Breeding of mammals in Rajasthan desert. *J Mamm* 41 386-389
5. Prakash I 1963 Zoogeography and evolution of the mammalian fauna of Rajasthan desert *Mammalia* 27 342-351
6. Prakash I 1974 The ecology of vertebrates of the Indian desert in *Biogeography and Ecology in India* Ed. M S Mani pp 369-420 (The Hague : Dr Junk bv)
7. Prakash I 1975 The ecology and zoogeography of mammals in *Environmental Analysis of the Thar Desert* Eds. R K Gupta and Ishwar Prakash pp 448-467 (Dehradun : English Book Depot)



8. Prakash Ishwar 1977. *The Amazing Life in the Indian Desert* pp. 1-18 (Jodhpur : Central Arid Zone Research Institute Monograph 6)
9. Prakash Ishwar 1986 Faunal diversity in the arid tracts *Proc. Indian Acad Sci Sp.* No. 45-58
10. Prakash Ishwar 1988 Wildlife - human - animal interactions, and conservation in the Rajasthan desert in *Desert Ecology* Ed. Ishwar Prakash pp. 169-178 (Jodhpur: Scientific Publishers)
11. Prakash Ishwar 1988a Changing patterns of rodent populations in India in *Rodent Pest Management* Ed. Ishwar Prakash pp 179-190 (Boca Raton, USA: CRC Press)
12. Prakash Ishwar 1989 Degradation of desert rangeland due to grazing pressure in *Perspectives in Ecology* Eds J S Singh and Brij Gopal pp. 181-191 (New Delhi : Jagmandir Book Agency)
14. Prakash Ishwar 1992 Conservation of biological resources in the risk prone Indira Gandhi Nahar Command area in the Thar desert. *Proc Natl Workshop "Sustainable Development of High Risk Environment"* Ed. Anil Gupta *In press*
15. Prakash Ishwar Ghosh P K 1980 Human animal interactions in the Rajasthan desert. *J Bombay nat Hist Soc* 75 1259-1261
16. Prakash Ishwar and Ghosh P K 1992 *Rodents in Indian Agriculture* pp 1-707 (Jodhpur: Scientific Publishers)
17. Prakash Ishwar and Mathur R P 1987 *Management of Rodent Pests* pp. 1-133 (New Delhi: Indian Council of Agricultural Research)
18. Saxena, S K and Prakash, Ishwar 1992 Rehabilitation of arid grazing lands in the Thar desert; in *Ecosystem Rehabilitation* Ed. M K Wali Vol. 2: 37-50 (The Hague: SPB Academic Press bv).

## AUTHOR INDEX

|                     |      |                 |      |
|---------------------|------|-----------------|------|
| Antia N H           | 1596 | Mitra A P       | 1023 |
| Appaji Rao N        | 1337 | Mohan Ram H Y   | 1439 |
| Awasthi D D         | 1790 | Mukherjee S K   | 1401 |
| Bachhawat B K       | 1085 | Mukunda N       | 1269 |
| Bachhawat B K       | 1707 | Nair N B K      | 1465 |
| Bamji Mahtab S      | 1720 | Notani N K      | 1732 |
| Bhargava P M        | 1741 | Pant D D        | 1846 |
| Bimla Buti          | 1286 | Prem Narain     | 1908 |
| Chandrasekhar S     | 1196 | Puri V          | 1761 |
| Chandrasekharan M K | 1255 | Rajaraman R     | 1932 |
| Dass C M S          | 1402 | Ramamurti B     | 1614 |
| Desikachari T V     | 1747 | Ramanna R       | 1100 |
| Dinesh Mohan        | 1066 | Rao C N R       | 1178 |
| Gopalan C           | 999  | Reddy A K N     | 1527 |
| Gupta C M           | 1684 | Roonwal M L     | 1943 |
| Indira Nath         | 1672 | Sasisekharan V  | 1321 |
| Ishwar Prakash      | 1984 | Seal S C        | 1628 |
| Jauhari B M         | 1811 | Seshadri C V    | 1545 |
| Jha S S             | 1237 | Sethi P K       | 1567 |
| Kalbag S S          | 1560 | Siddiqi O       | 1228 |
| Kannan K K          | 1302 | Singh B N       | 1874 |
| Kanungo M S         | 1423 | Srivastava B S  | 1698 |
| Khoshoo T N         | 1895 | Srivastava P N  | 1215 |
| Kurien V            | 1584 | Subramaniam C V | 1828 |
| Matra P K           | 1353 | Sukh Dev        | 1140 |
| Mehra P N           | 1362 | Tandon P N      | 1499 |
| Mishra R            | 1804 | Valiathan M V S | 1519 |

